

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
22 September 2005 (22.09.2005)

PCT

(10) International Publication Number
WO 2005/086891 A3

- (51) International Patent Classification⁷: **C12Q 1/68**
- (21) International Application Number:
PCT/US2005/007894
- (22) International Filing Date: 7 March 2005 (07.03.2005)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/550,810 5 March 2004 (05.03.2004) US
60/604,076 24 August 2004 (24.08.2004) US
60/650,401 4 February 2005 (04.02.2005) US
- (71) Applicants (*for all designated States except US*):
ROSETTA INPHARMATICS LLC [US/US]; 401 Terry Avenue North, Seattle, WA 98109 (US). **THE NETHERLANDS CANCER INSTITUTE** [NL/NL]; Plesmanlaan 121, NL-1066 CX Amsterdam (NL).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **DAI, Hongyue** [CN/US]; 16814 118th Avenue NE, Bothell, WA 98011 (US). **VAN'T VEER, Laura, J.** [NL/NL]; Brouwersgracht 192-G, NL-1013 HC Amsterdam (NL). **LAMB, John** [GB/US]; 1216 N 172nd Street, Shoreline, WA 98133 (US). **STOUGHTON, Roland** [US/US]; Apt. D, 5919 Mildred Street, San Diego, CA 92110 (US). **FRIEND, Stephen, H.** [US/US]; 101 W. Mermaid Lane, Philadelphia, PA 19118 (US). **HE, Yudong** [US/US]; 11410 NE 124th Street #148, Kirkland, WA 98034 (US).
- (74) Agents: **ANTLER, Adriane, M.** et al.; Jones Day, 222 East 41st Street, New York, NY 10017-6702 (US).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— *with international search report*
- (88) Date of publication of the international search report:
9 March 2006
- (48) Date of publication of this corrected version:
4 May 2006
- (15) Information about Corrections:
see PCT Gazette No. 18/2006 of 4 May 2006
Previous Correction:
see PCT Gazette No. 04/2006 of 26 January 2006
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: CLASSIFICATION OF BREAST CANCER PATIENTS USING A COMBINATION OF CLINICAL CRITERIA AND INFORMATIVE GENESETS

(57) Abstract: The present invention provides prognostic methods for conditions such as cancer, for example, breast cancer, comprising classifying an individual by a plurality of phenotypic, genotypic or clinical characteristics of the condition into a plurality of patient subsets, and analyzing the pattern of expression of prognosis-informative genes identified for that subset in a sample from the individual. The present invention also provides methods for constructing such patient subsets and of identifying prognosis-informative genesets for such subsets. The invention further provides methods of assigning a therapeutic regimen to an individual, microarrays useful for performing prognosis, kits comprising these microarrays, and computer systems and programs for implementing the methods of the invention.



WO 2005/086891 A3

CLASSIFICATION OF BREAST CANCER PATIENTS USING A COMBINATION OF CLINICAL CRITERIA AND INFORMATIVE GENESETS

[0001] This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Patent Application No. 60/650,401, filed on February 4, 2005, U.S. Provisional Patent Application No. 60/604,076, filed on August 24, 2004, and U.S. Provisional Patent Application No. 60/550,810, filed on March 5, 2004, each of which is incorporated by reference herein in its entirety.

1. FIELD OF THE INVENTION

[0002] The present invention relates to the use of both phenotypic and genotypic aspects of a condition, such as a disease, in order to identify discrete subsets of patients for which specific sets of informative genes are then identified. The invention also relates to the classification of individuals, such as breast cancer patients, into a subset of the condition on the basis of clinical parameters and the status of markers, for example, of genes expression patterns, and the prognosis of those individuals on the basis of markers informative for prognosis within the subset of the condition. The invention also relates to methods of determining a course of treatment or therapy to an individual having, or suspected of having, a condition, such as breast cancer. The invention further relates to methods of structuring a clinical trial, particularly using five breast cancer-specific patient subsets and prognosis-informative genes for each, and of identifying patient populations for clinical trials or for other condition-related, for example, breast cancer-related, research. Finally, the invention relates to computer implementations of the above methods.

2. BACKGROUND OF THE INVENTION

[0003] The increased number of cancer cases reported in the United States, and, indeed, around the world, is a major concern. Currently there are only a handful of treatments available for specific types of cancer, and these provide no guarantee of success. In order to be most effective, these treatments require not only an early detection of the malignancy, but a reliable assessment of the severity of the malignancy.

[0004] The incidence of breast cancer, a leading cause of death in women, has been gradually increasing in the United States over the last thirty years. Its cumulative risk is relatively high; 1 in 8 women are expected to develop some type of breast cancer by age 85 in the United States. In fact, breast cancer is the most common cancer in women and the second

most common cause of cancer death in the United States. In 1997, it was estimated that 181,000 new cases were reported in the U.S., and that 44,000 people would die of breast cancer (Parker *et al.*, *CA Cancer J. Clin.* 47:5-27 (1997); Chu *et al.*, *J. Nat. Cancer Inst.* 88:1571-1579 (1996)). While mechanism of tumorigenesis for most breast carcinomas is largely unknown, there are genetic factors that can predispose some women to developing breast cancer (Miki *et al.*, *Science*, 266:66-71(1994)).

[0005] Sporadic tumors, those not currently associated with a known germline mutation, constitute the majority of breast cancers. It is also likely that other, non-genetic factors also have a significant effect on the etiology of the disease. Regardless of the cancer's origin, breast cancer morbidity and mortality increases significantly if it is not detected early in its progression. Thus, considerable effort has focused on the early detection of cellular transformation and tumor formation in breast tissue.

[0006] A marker-based approach to tumor identification and characterization promises improved diagnostic and prognostic reliability. Typically, the diagnosis of breast cancer requires histopathological proof of the presence of the tumor. In addition to diagnosis, histopathological examinations also provide information about prognosis and selection of treatment regimens. Prognosis may also be established based upon clinical parameters such as tumor size, tumor grade, the age of the patient, and lymph node metastasis.

[0007] Diagnosis and/or prognosis may be determined to varying degrees of effectiveness by direct examination of the outside of the breast, or through mammography or other X-ray imaging methods (Jatoi, *Am. J. Surg.* 177:518-524 (1999)). The latter approach is not without considerable cost, however. Every time a mammogram is taken, the patient incurs a small risk of having a breast tumor induced by the ionizing properties of the radiation used during the test. In addition, the process is expensive and the subjective interpretations of a technician can lead to imprecision. For example, one study showed major clinical disagreements for about one-third of a set of mammograms that were interpreted individually by a surveyed group of radiologists. Moreover, many women find that undergoing a mammogram is a painful experience. Accordingly, the National Cancer Institute has not recommended mammograms for women under fifty years of age, since this group is not as likely to develop breast cancers as are older women. It is compelling to note, however, that while only about 22% of breast cancers occur in women under fifty, data suggests that breast cancer is more aggressive in pre-menopausal women.

[0008] In clinical practice, accurate diagnosis of various subtypes of breast cancer is important because treatment options, prognosis, and the likelihood of therapeutic response all

vary broadly depending on the diagnosis. Accurate prognosis, or determination of distant metastasis-free survival could allow the oncologist to tailor the administration of adjuvant chemotherapy, with women having poorer prognoses being given the most aggressive treatment. Furthermore, accurate prediction of poor prognosis would greatly impact clinical trials for new breast cancer therapies, because potential study patients could then be stratified according to prognosis. Trials could then be limited to patients having poor prognosis, in turn making it easier to discern if an experimental therapy is efficacious.

[0009] To date, no set of satisfactory predictors for prognosis based on the clinical information alone has been identified. Many have observed that the ER status has a dominant signature in the breast tumor gene expression profiling. See West *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, *Nature* 415:530 (2002); Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001); Sotiriou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:10393 (2003). It is generally accepted that there is some relationship between patient survival and ER status. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002); Surowiak *et al.*, *Folia Histochem. Cytobiol.* 39:143 (2001); Pichon *et al.*, *Br. J. Cancer* 73:1545 (1996); Collett *et al.*, *J. Clin. Pathol.* 49:920 (1996). *BRCA1* mutations are related to the familial cancer susceptibility. Biesecker *et al.*, *JAMA* 269:1970 (1993); Easton *et al.*, *Cancer Surv.* 18:95 (1993). Age is also considered to be a prognosis factor since young cancer patients tend to have poor tumors. Maggard *et al.*, *J. Surg. Res.* 113:109 (2003). Lymph node status is a factor in deciding the treatment. Eifel *et al.*, *J. Natl. Cancer Inst.* 93:979 (2001).

[0010] The discovery and characterization of *BRCA1* and *BRCA2* has recently expanded our knowledge of genetic factors which can contribute to familial breast cancer. Germ-line mutations within these two loci are associated with a 50 to 85% lifetime risk of breast and/or ovarian cancer (Casey, *Curr. Opin. Oncol.* 9:88-93 (1997); Marcus *et al.*, *Cancer* 77:697-709 (1996)). Only about 5% to 10% of breast cancers, however, are associated with breast cancer susceptibility genes, *BRCA1* and *BRCA2*. The cumulative lifetime risk of breast cancer for women who carry the mutant *BRCA1* is predicted to be approximately 92%, while the cumulative lifetime risk for the non-carrier majority is estimated to be approximately 10%. *BRCA1* is a tumor suppressor gene that is involved in DNA repair and cell cycle control, which are both important for the maintenance of genomic stability. More than 90% of all mutations reported so far result in a premature truncation of the protein product with abnormal or abolished function. The histology of breast cancer in *BRCA1* mutation carriers differs from that in sporadic cases, but mutation analysis is the only way to find the carrier.

Like *BRCA1*, *BRCA2* is involved in the development of breast cancer, and like *BRCA1* plays a role in DNA repair. However, unlike *BRCA1*, it is not involved in ovarian cancer.

[0011] Other genes have been linked to breast cancer, for example c-erb-2 (*HER2*) and p53 (Beenken *et al.*, *Ann. Surg.* 233(5):630-638 (2001). Overexpression of c-erb-2 (*HER2*) and p53 have been correlated with poor prognosis (Rudolph *et al.*, *Hum. Pathol.* 32(3):311-319 (2001), as has been aberrant expression products of *mdm2* (Lukas *et al.*, *Cancer Res.* 61(7):3212-3219 (2001) and cyclin1 and p27 (Porter & Roberts, International Publication WO98/33450, published August 6, 1998).

[0012] The detection of *BRCA1* or *BRCA2* mutations represents a step towards the design of therapies to better control and prevent the appearance of these tumors. Recently, many studies have used gene expression profiling to analyze various cancers, and those studies have provided new diagnosis and prognosis information in the molecular level. See Zajchowski *et al.*, "Identification of Gene Expression Profiles that Predict the Aggressive Behavior of Breast Cancer Cells," *Cancer Res.* 61:5168 (2001); West *et al.*, "Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles," *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, "Gene Expression Profiling Predicts the Outcome of Breast Cancer," *Nature* 415:530 (2002); Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Khan *et al.*, *Cancer Res* 58, 5009 (1998); Golub *et al.*, *Science* 286, 531 (1999); DeRisi *et al.*, *Nat. Genet.* 14:457 (1996); Alizadeh *et al.*, *Nature* 403, 503 (2000). Methods for the identification of informative genesets for various cancers have also been described. See Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Golub *et al.*, United States Patent No. 6,647,341.

[0013] Genesets have been identified that are informative for differentiating individuals having, or suspected of having, breast cancer based on estrogen receptor (ER) status, or *BRCA1* mutation vs. sporadic (*i.e.*, other than *BRCA1*-type) mutation status. See Roberts *et al.*, WO 02/103320; van't Veer *et al.*, *Nature* 415:530 (2001). Genesets have also been identified that enable the classification of sporadic tumor-type individuals as those who will likely have no metastases within five years of initial diagnosis (*i.e.*, individuals with a good prognosis) or those who will likely have a metastasis within five years of initial diagnosis (*i.e.*, those having a poor prognosis). Roberts, *supra*; van't Veer, *supra*.

[0014] Roberts *et al.* WO 02/103320 describes a 70-gene set, useful for the prognosis of breast cancer, which outperformed clinical measures of prognosis, and which showed good

potential in selecting good outcome patients, thereby avoiding over-treatment. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002). The expression of genes with most predictive value, however, were not homogeneous among poor patients, suggesting the need for improvement.

[0015] Although the patterns of gene expression as described in Roberts *et al.* were correlated with existing clinical indicators such as estrogen receptor and *BRCA1* status, clinical measures were not incorporated. Furthermore, although the poor-outcome group in particular showed heterogeneity in expression pattern, the best classifier decision rule found during these studies was a fairly simple one based on the similarity of a patient profile to the average profile of a good-outcome training group.

[0016] It is evident that breast cancer is the result of more than one type of molecular event. Likewise, a variety of other conditions, such as other cancers; non-cancer diseases such as diabetes, autoimmune or neurodegenerative disorders, obesity; etc., are also the result of more than one molecular event. Moreover, an individual's response to exposure to particular environmental conditions, for example, exposure to natural or man-made agents, such as toxins, pollutants, drugs, food additives, etc., likely result from more than one molecular event. Thus, there exists a need for improved prognostic methods so that appropriate courses of prophylaxis and/or therapy may be provided. Genesets having improved prognostic power can be identified by first identifying discrete subsets of individuals based on genotypic or phenotypic characteristics relevant to the disease or condition, and then identifying genesets informative for prognosis within those subsets of patients. Individuals having the condition, or who are suspected of having the condition, such as breast cancer, would then be provided therapies appropriate to the molecular mechanisms underlying the condition. The present invention provides such methods for breast cancer, and for other cancers, diseases or conditions.

3. SUMMARY OF THE INVENTION

[0017] The present invention provides methods of identifying relevant subsets of conditions, and the identification of markers relevant to those subsets, for example, for prognosis of individuals classifiable into one of those subsets. The invention further provides sets of markers useful for the prognosis of individuals having breast cancer, wherein those patients have been classified according to one or more characteristics of breast cancer.

[0018] Thus, the present invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the

basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; and (b) identifying within each of said first classes a first set of genes or markers informative for said condition, wherein said first set of genes or markers within each of said first classes is unique to said class relative to other first classes. In a specific embodiment, this method further comprises additionally classifying into a plurality of second classes said samples or individuals in at least one of said first classes on the basis of a phenotypic or genotypic characteristic different than that used in said classifying step (a); and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other first and second classes.

[0019] The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes; (b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used in said classifying step (a); and (c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition, wherein said second set of genes or markers is unique to said class relative to other first and second classes.

[0020] The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics; (b) identifying at least two first condition classes differentiable by said first characteristic; (c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and (d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

[0021] The invention further provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient

class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression. In a specific embodiment, said condition is cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In a more specific embodiment, said cancer is breast cancer. In another specific embodiment, said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a poor prognosis. In a more specific embodiment, said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a good prognosis. In a more specific embodiment, said classifying in step (c) is carried out by a method comprising comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said plurality of patient classes comprises ER^{-} , *BRCA1* individuals; ER^{-} , sporadic individuals; ER^{+} , *ER/AGE* high individuals; ER^{+} , *ER/AGE* low, *LN*⁺ individuals; and ER^{+} , *ER/AGE* low, LN^{-} individuals.

[0022] The invention further provides a method of classifying a breast cancer patient as having a good prognosis or a poor prognosis comprising: (a) classifying said breast cancer patient as ER^{-} , *BRCA1*; ER^{-} , sporadic; ER^{+} , *ER/AGE* high; ER^{+} , *ER/AGE* low, *LN*⁺; or ER^{+} , *ER/AGE* low, LN^{-} ; (b) determining the level of expression of a first plurality of genes in a cell sample taken from said breast cancer patient relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers in Table 1 if said breast cancer patient is classified as ER^{-} , *BRCA1*; in Table 2 if said breast cancer patient is classified as ER^{-} sporadic; in Table 3 if said breast cancer patient is classified as ER^{+} , *ER/AGE* high; in Table 4 if said breast cancer patient is classified as ER^{+} , *ER/AGE* low,

LN+; or in Table 5 if said breast cancer patient is classified as ER+, ER/AGE low, LN⁻; and (c) classifying said breast cancer patient as having a good prognosis or a poor prognosis on the basis of the level of expression of said first plurality of genes, wherein said breast cancer patient is “ER/AGE high” if the ratio of the $\log_{10}(\text{ratio})$ of ER gene expression to age exceeds a predetermined value, and “ER/AGE low” if the ratio of the $\log_{10}(\text{ratio})$ of ER gene expression to age does not exceed said predetermined value. In a specific embodiment, said control is the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, *BRCAl* individuals, if said breast cancer patient is ER⁻, *BRCAl*; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, sporadic individuals if said breast cancer patient is ER⁻, sporadic; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER+, ER/AGE high individuals, if said breast cancer patient is ER+, ER/AGE high; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER+, ER/AGE low, LN+ individuals where said breast cancer patient is ER+, ER/AGE low, LN+; or the average level of expression of each of said plurality of genes in a plurality of samples derived from ER+, ER/AGE low, LN⁻ individuals where said breast cancer patient is ER+, ER/AGE low, LN⁻. In a more specific embodiment, each of said individuals has a poor prognosis. In another more specific embodiment, each of said individuals has a good prognosis. In an even more specific embodiment, said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins in a sample from said breast cancer patient to said control, and classifying said breast cancer patient as having a poor prognosis if said level of expression correlates with said average level of expression of the corresponding genes or their encoded proteins in said control more strongly than would be expected by chance. In another specific embodiment, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises two of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises two of the genes for which markers are listed in Table 2. said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER+, ER/AGE high, and said plurality of genes comprises two of the genes for which markers are listed in Table 3. said

individual is ER+, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises two of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN-, and said plurality of genes comprises two of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER+, ER/AGE low, LN-, and said plurality of genes comprises all of the genes for which markers are listed in Table 4. In another specific embodiment, the method further comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[0023] In another embodiment, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising: (a) classifying said individual as ER-, *BRCAl*, ER-, sporadic; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE low, LN-; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER-, *BRCAl*; Table 2 if said individual is classified as ER-, sporadic; Table 3 if said individual is classified as ER+, ER/AGE high; Table 4 if said individual is classified as ER+, ER/AGE low, LN+; or Table 5 if said individual is classified as ER+, ER/AGE low, LN-; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another specific embodiment, said method further comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

[0024] The invention further provides a microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in any of Tables 1-5. The invention further provides a microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in Table 1, each of the genes for which markers are listed in Table 1, a plurality of the genes for which markers are listed in Table 2, each of the genes for which markers are listed in Table 2, a plurality of the genes for which markers are listed in Table 3, each of the genes for which markers are listed in Table 3, a plurality of the genes for which markers are listed in Table 4, each of the genes for which markers are listed in Table 4, a plurality of the genes for which markers are listed in Table 5, or each of the genes for which markers are listed in Table 5. The invention further provides any one of the above microarrays, wherein said probes are at least 50% of the probes on said microarray. The invention further provides any one of the above microarrays, wherein said probes are at least 90% of the probes on said microarray. The invention further provides microarray comprising probes complementary and hybridizable to a plurality of the genes for which markers are listed in any of Tables 1-5, wherein said probes are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and are complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

[0025] The invention further comprises a kit comprising any one of the above microarrays in a sealed container.

[0026] The invention further provides a method of identifying a set of genes informative for a condition, said condition having a plurality of phenotypic or genotypic characteristics such that samples may be categorized by at least one of said phenotypic or genotypic characteristics into at least one characteristic class, said method comprising: (a) selecting a plurality of samples from individuals having said condition; (b) identifying a first set of genes informative for said characteristic class using said plurality of samples; (c) predicting the characteristic class of each of said plurality of samples; (d) discarding samples for which said characteristic class is incorrectly predicted; (e) repeating steps (c) and (d) at least once; and (f) identifying a second set of genes informative for said characteristic class using samples in said plurality of samples remaining after step (e).

[0027] The invention further provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising: (a) classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; (b) determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category; (c) determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and (d) assigning the individual to a category in a clinical trial on the basis of prognosis.

[0028] The invention also provides a method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising: (a) subdividing a plurality of individuals or samples derived therefrom of the organism subject to the condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of the organism, wherein each class consists of a plurality of individuals or samples derived therefrom of the organism each of which having one or more clinical, phenotypic or genotypic characteristics specific for the class; and (b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein, if a set of genes or markers informative for said condition in individuals in said class is obtained for any of said one or more of said plurality of classes, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

[0029] In one embodiment, the method further comprises, for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of individuals or samples derived therefrom in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein for each of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

[0030] The invention also provides a method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising: (a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of

said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class; (b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein if a set of genes or markers informative for said condition in individuals in said class is identified for any of said one or more of said classes, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism; and (c) for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of samples or individuals in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used those used for defining said class, wherein for each of one or more of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism.

[0031] In the methods of the invention, the condition can be a type of cancer. In such an embodiment, each of said sets of genes or markers can be informative of prognosis of individuals in a corresponding class. In one embodiment, the condition is breast cancer, and the one or more clinical, phenotypic or genotypic characteristics comprise age, ER level, ER/AGE, BRAC1 status, and lymph node status.

[0032] In one embodiment, the methods of the invention further comprise generating a template profile comprising measurements of levels of genes or markers of the set of informative genes or markers for said class representative of levels of the genes or markers in a plurality of patients having a chosen prognosis level.

[0033] The invention also provides a method for predicting a breast cancer patient as having a good prognosis or a poor prognosis, comprising: (a) classifying said breast cancer patient into one of the following classes: (a1) ER⁻, *BRCAI*; (a2) ER⁻, sporadic; (a3) ER⁺, ER/AGE high; (a4) ER⁺, ER/AGE low, LN⁺; or (a5) ER⁺, ER/AGE low, LN⁻; (b) determining a profile comprising measurements of a plurality of genes or markers in a cell sample taken from said breast cancer patient, said plurality of genes markers comprising at least two of the genes or markers corresponding to the markers in (b1) Table 1 if said breast cancer patient is classified as ER⁻, *BRCAI*; (b2) Table 2 if said breast cancer patient is classified as ER⁻ sporadic; (b3) Table 3 if said breast cancer patient is classified as ER⁺, ER/AGE high; (b4)

Table 4 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁺; or (b5) Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and (c) classifying said breast cancer patient as having a good prognosis or a poor prognosis based on said profile of said plurality of genes or markers, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, wherein said ER/AGE is a metric of said ER level relative to the age of said patient, and wherein LN⁺ designates a greater than 0 lymph nodes status in said patient and LN⁻ designates a 0 lymph nodes status in said patient.

[0034] In one embodiment, step (c) is carried out by a method comprising comparing said profile to a good prognosis template and/or a poor prognosis template, and wherein said patient is classified as having a good prognosis if said profile has a high similarity to a good prognosis template or has a low similarity to a poor prognosis template or as having a poor prognosis if said profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template. A good prognosis template comprises measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of good outcome patients, while a poor prognosis template comprises measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of poor outcome patients. Here a good outcome patient is a breast cancer patient who has non-reoccurrence of metastases within a first period of time after initial diagnosis, while a poor outcome patient is a patient who has reoccurrence of metastases within a second period of time after initial diagnosis.

[0035] In another embodiment, the methods for predicting the prognosis of a breast cancer patient further comprise determining said profile, said ER level, said LN status, and/or, said ER/AGE. In one embodiment, said profile is an expression profile comprising measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

[0036] In one embodiment, said expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

[0037] In one embodiment, the measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

[0038] In one embodiment, the similarity of said expression profile to said good or poor prognosis template is represented by a correlation coefficient between said expression profile and said good or poor prognosis template, respectively, and a correlation coefficient greater than a correlation threshold, e.g., 0.5, indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

[0039] In another embodiment, the similarity of said expression profile to said good or poor prognosis template is represented by a distance between said cellular constituent profile and said good or poor prognosis template, respectively, and a distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

[0040] In another embodiment, said profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

[0041] In one embodiment, said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor, e.g., the estrogen receptor α gene, in said patient relative to expression level of said gene in said control sample, and said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

[0042] In one embodiment, said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

[0043] In a specific embodiment, said estrogen receptor level is measured by a polynucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, said control sample is a pool of breast cancer cells of different patients, and $c = 0.1$ and $d = 42.5$.

[0044] In one embodiment, said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients. In another embodiment, said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

[0045] In one embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 1. In one embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

[0046] In another embodiment, the individual is ER⁻, sporadic, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 2. In one embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

[0047] In still another embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 3. In one embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

[0048] In still another embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4. In one embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

[0049] In still another embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4. In one embodiment, the individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

[0050] In one embodiment, said profile further comprises one or more genes for which markers are not found in Tables 1-5, which are informative for prognosis.

[0051] The invention also provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising assigning said individual to one category in a clinical trial if said individual has a good prognosis as determined by any one of the methods described above, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis as determined by any one of the methods described above.

[0052] In one embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual based on said profile, said ER level, said LN status, and/or, said ER/AGE.

[0053] In one embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of one or more other clinical, phenotypic or genotypic characteristic of breast cancer.

[0054] In one embodiment, the method further comprises determining in said cell sample the levels of expression of said one or more genes for which markers are not found in Tables 1-5, and determining from said expression levels of said one or more genes, whether said individual has a good prognosis or a poor prognosis.

4. BRIEF DESCRIPTION OF THE DRAWINGS

[0055] FIG. 1 depicts the decision tree that resulted in the five patient subsets used to identify informative prognosis-related genes.

[0056] FIG. 2: Relationship between ER level and age. (A) Scatter plot of ER vs. age for ER+ patients. Black dots indicate metastases free samples, and gray dots indicate metastases samples. It appears that patients of ER+ group can be subdivided into “ER+, ER/AGE high” group (above the black line) and “ER+, ER/AGE low” (below the black line) group. The black line is approximated by $ER = 0.1 * (AGE - 42.5)$, and the dashed line by $ER = 0.1 * (age - 50)$. Within each population, the ER level also increases with age. (B) Age distribution of all patients in ER+ samples. A bimodal distribution is observed. (C) ER-modulated age (age – 10*) distribution of all patients in ER+ samples. A bimodal distribution is observed. (D) Age distribution of samples with metastasis. (E) ER-modulated age distribution of samples with metastasis. The three peaks appearing in this distribution suggest a polymorphism.

[0057] FIG. 3. Performance of classifier for the “ER-/sporadic” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0058] FIG. 4. Performance of classifier for the “ER+, ER/AGE high” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0059] FIG. 5. Performance of classifier for the “ER+, ER/AGE low/LN⁻” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicates samples with metastases. Dashed line indicates the threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0060] FIG. 6. Performance of classifier for the “ER+, ER/AGE low/LN+” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0061] FIG. 7. Performance of classifier for the “ER⁻, *BRCAl*” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group

(poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

[0062] FIG. 8. Heatmaps of genes representing key biological functions in subgroups of patients: A: Cell cycle genes are predictive of outcome in patients with ER/age high. B: Cell cycle genes are not predictive of outcome in “ER- and sporadic” patients C: Glycolysis genes are predictive of outcome in patients with ER/age low and LN-. D: Glycolysis genes are not predictive of outcome in “ER- & BRCA1” patients.

5. DETAILED DESCRIPTION OF THE INVENTION

5.1 INTRODUCTION

[0063] The present invention provides methods for classifying individuals having a condition, such as a disease, into one or more subsets of individuals, where individuals in each subset are characterized by one or more phenotypic or genotypic characteristics of the condition. The individuals may be eukaryotes or prokaryotes, may be animals such as mammals, including but not limited to humans, primates, rodents, felines, canines, etc., birds, reptiles, fish, etc. “Individuals” as used herein also encompasses single-celled organisms, or colonies thereof, such as bacteria and yeast. The condition may be a disease, such as cancer, and may be a specific cancer, such as breast cancer. The condition may also be an environmental condition, such as exposure to a toxin, pollutant, drug, proximity to urban or industrial areas, etc.

[0064] The present invention provides methods of determining the prognosis of individuals having a condition, such as cancer, for example, breast cancer, or who are suspected of having the condition, by the use of a combination of clinical, biological or biochemical parameters of the condition and gene expression pattern data. For prognosis, the parameters selected preferably relate to or affect the progression and/or outcome of the condition. The pattern of gene expression within a subset of individuals having the particular condition leads to the identification of sets of genes within a subset that is informative for that subset, for example, for prognosis within that subset. In general, the successful identification of sets of genes informative for prognosis within a particular subset justifies the selection of the plurality of clinical, biological or biochemical parameters of the condition on which division of individuals into condition subsets is based.

[0065] In the example of breast cancer, patient groups are first classified according to at least one of age, lymph node (LN) status, estrogen receptor (ER) level, and *BRCA1* mutation status into discrete patient subsets. These clinical factors have been implicated in tumor

etiology as well as differences in disease outcome. These characteristics are not limiting; other genotypic or phenotypic characteristics of breast cancer, for example, tumor grade, tumor size, tumor cell type, etc., may also be used, alone or in combination with those listed herein, in order to classify individuals. The differences in gene expression or in tumor fate related to these parameters likely represent differences in tumor origin and tumor genesis, and are therefore good candidates for tumor stratification. Genesets informative for prognosis within each subset are then identified. New breast cancer patients are then classified using the same criteria, and a prognosis is made based on the geneset specific for the patient subset into which the patient falls. In the process of constructing a prognosis classifier within each patient subset, particular attention is paid to the homogeneous patterns related to the tumor outcome. Emergence of such homogeneous prognosis patterns may indicate the most common mechanism to metastasis within a subset. At the same time, successful identification of such patterns also justifies the parameters being used for the tumor stratification. To differentiate this approach from an mRNA-alone approach, the current approach of integrating clinical data with the gene expression data is referred to herein as a “comprehensive prognosis”.

5.2 DEFINITIONS

[0066] As used herein, “*BRCA1* tumor” or “*BRCA1* type” means a tumor having cells containing a mutation of the *BRCA1* locus.

[0067] The “absolute amplitude” of correlation means the absolute value of the correlation; e.g., both correlation coefficients -0.35 and 0.35 have an absolute amplitude of 0.35.

[0068] “Marker” means a cellular constituent, or a modification of a cellular constituent (e.g., an entire gene, EST derived from that gene, a protein encoded by that gene, post-translational modification of the protein, etc.) the expression or level of which changes between certain conditions. Where a change in a characteristic of the constituent correlates with a certain condition, the constituent is a marker for that condition.

[0069] “Marker-derived polynucleotides” means the RNA transcribed from a marker gene, any cDNA or cRNA produced therefrom, and any nucleic acid derived therefrom, such as synthetic nucleic acid having a sequence derived from the gene corresponding to the marker gene.

[0070] A “similarity value” is a number that represents the degree of similarity between two things being compared. For example, a similarity value may be a number that indicates the overall similarity between a patient’s expression profile of specific phenotype-related

markers and a template specific to that phenotype (for instance, the similarity to a “good prognosis” template, where the phenotype is a good prognosis). The similarity value may be expressed as a similarity metric, such as a correlation coefficient, or may simply be expressed as the expression level difference, or the aggregate of the expression level differences, between a patient sample and a template.

[0071] A “patient subset” is a group of individuals, all of whom have a particular condition, or are subject to a particular condition, which is distinguished from other individuals having that condition by one or more phenotypic, genotypic or clinical characteristics of the condition, or of a response to the condition. For example, where the condition is breast cancer, individuals may belong to an “ER⁺” or an “ER⁻” patient subset, or may belong to a particular age group patient subset.

[0072] A gene and/or marker is “informative” for a condition, phenotype, genotype or clinical characteristic if the expression of the gene or marker is correlated or anticorrelated with the condition, phenotype, genotype or clinical characteristic to a greater degree than would be expected by chance.

[0073] An individual of a given age can be classified as “ER/AGE high” if the individual’s ER level is higher than a threshold value for the given age. The threshold can be age-dependent, i.e., a different threshold for each different age. In one embodiment, the age-dependent threshold value is calculated as $c \cdot (AGE - d)$, where c is a coefficient, AGE is the age of the patient, and d is an age threshold. The parameters c and d depend on the ER level and AGE used. They can be determined by fitting patients’ ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a specific embodiment, $c = 0.1$ and $d = 42.5$ is used for ER levels represented by a log(ratio) of ER expression level. Thus, for example, the threshold for a 45-year old individual in this embodiment is $0.1 (45 - 42.5)$, or 0.25, and if the log(ratio) of ER expression level of the individual is equal to or greater than 0.25, the individual is classified as “ER/AGE high”; otherwise, the individual is classified as “ER/AGE low.”

5.3 IDENTIFICATION OF DIAGNOSTIC AND PROGNOSTIC MARKER SETS

5.3.1 IDENTIFICATION OF CONDITION SUBSETS

[0074] The present invention provides methods of identifying sets of genes and/or markers useful in the diagnosis and prognosis of breast cancer. More generally, the invention also provides methods of identifying sets of genes and/or markers useful in the diagnosis or prognosis of other cancers, and even more generally, of identifying sets of genes and/or

markers useful in the differentiation between subgroups of individuals having a particular condition, such as a disease or exposure to a particular environmental condition.

[0075] The method may be applied to any condition for which a plurality of phenotypic or genotypic subsets may be identified. The condition may be a disease; for example, the condition may be cancer, an autoimmune disease, an inflammatory disease, an infectious disease, a neurological disease, a degenerative disease, etc. The condition may be environmental; for example, the condition may be a particular diet, geographic location, etc.; the condition may be exposure to a compound, including, for example, a drug, a toxin, a carcinogen, a foodstuff, a poison, an inhaled compound, an ingested compound, etc.; the condition may be a particular genetic background or predisposition to a medical condition; etc.

[0076] Where the condition is cancer, the condition may be any cancer, for example, without limitation: leukemias, including acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic leukemia, promyelocytic leukemia, myelomonocytic leukemia, monocytic leukemia, and erythroleukemia; chronic leukemia, such as chronic myelocytic (granulocytic) leukemia or chronic lymphocytic leukemia; polycythemia vera; lymphomas, such as Hodgkin's disease and non-Hodgkin's disease; multiple myeloma; Waldenström's macroglobulinemia; heavy chain disease; solid tumors, such as sarcomas and carcinomas, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilms' tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, meningioma, melanoma, neuroblastoma, or retinoblastoma; etc.

[0077] Rather than stratifying individuals, such as patients or tumor samples derived from patients, by gene expression patterns in the first instance, the method of identifying sets of genes informative for a condition begins by identifying phenotypic, genotypic or clinical

subsets of individuals within the larger class of individuals having or affected by the condition.

[0078] In one embodiment, the condition is cancer, and the subsets are distinguished by phenotypic, genotypic, and/or clinical characteristics of the cancer. In this embodiment, groups of individuals are classified according to one or more phenotypic, genotypic, or clinical characteristics relevant to the cancer into patient subsets. At any step in the process of subdividing a patient population into patient subsets, the expression level of one or more genes may be determined in order to identify whether a prognosis-informative set of genes may be identified for the particular patient subset. If an informative gene set is identified, but is not as informative as desired, the patient subset may be further divided and a new geneset identified. These subsets may be further subdivided. For example, a group of individuals affected by a particular cancer may be classified first on the basis of a phenotypic, genotypic or clinical characteristic A into subsets S1 and S2. The levels of expression of a plurality of genes are then determined in tumor samples taken from individuals that fall within subsets S1 or S2 in order to identify sets of genes informative for prognosis within these subsets. Subsets S1 and S2 may then each be subdivided into two or more subsets based on other phenotypic, genotypic or clinical characteristics. The basis for subdivision, if performed, need not be the same for S1 and S2. For example, in various embodiments, S1 is not subdivided, while S2 is subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B while S2 is not subdivided; or S1 and S2 are both subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B, while S2 is subdivided according to characteristic C; and so on. For a particular decision matrix leading to a plurality of patient subsets, the preferred outcome is a prognosis-informative set of genes for each patient subset. Different decision matrices may lead to different patient subsets, which, in turn, may result in different sets of prognosis-informative genes.

[0079] In the specific example of breast cancer, a plurality of phenotypic, genotypic or clinical indications are used to classify a patient as being a member of one of a plurality of patient subsets, wherein the indications are medically, biochemically or genetically relevant to breast cancer. For example, a group of patients may be classified into patient subsets based on criteria including, but not limited to, estrogen receptor (ER) status, type of tumor (*i.e.*, *BRCA1*-type or sporadic), lymph node status, grade of cancer, invasiveness of the tumor, or age. “*BRCA1*-type” indicates that the *BRCA1* mutation is present. In each classification step, a group of cancer patients may be classified into only two classes, for example, ER⁺ or ER⁻, or into three or more subsets (for example, by tumor grade), depending upon the

characteristic used to determine the subsets. As used herein, “ER+” indicates that the estrogen receptor is expressed at some elevated level; for example, it may indicate that the estrogen receptor is detectably expressed, or may indicate that more than 10% of cells are histologically stained for the receptor, etc. Conversely, “ER-” indicates that the estrogen receptor is expressed at a reduced level or not at all; for example, it may indicate that the receptor is not detectably expressed, or that 10% or less of cells are histologically stained for the receptor, etc. Marker gene sets optimized for each phenotypic class are preferably determined after the subsets are established. Where informative markers for a particular patient subset, distinguished from another subset by a particular characteristic of the condition of interest, cannot be determined, the subset may be further divided by another characteristic of the condition to create a plurality of second patient subsets, whereupon genes informative for these second patient subsets may be identified.

[0080] FIG. 1 depicts the process, described in the Examples, of subdivision of a collection of breast cancer patients according to phenotypic and genotypic characteristics relevant to breast cancer, in preparation for identification of genes informative for prognosis. A collection of breast cancer tumor samples was first subdivided by estrogen receptor status. ER status was chosen because the presence or absence of the estrogen receptor greatly influences the expression of other genes. In the ER+ patient subset, it was noted that patients appeared to be bimodally distributed by ER level vs. age; that is, ER level dependence upon age tended to fall within two classes, as separated by the solid line in FIG. 2A. This bimodality was used to further subdivide ER+ individuals into “ER+, ER/AGE high” individuals and “ER+, ER/AGE low” individuals. A set of informative genes was identified for the ER+, ER/AGE high patient subset. An informative set was not identified for the ER+, ER/AGE low subset, however, so the subset of patients was further divided into LN+ and LN- individuals. Thus, in one embodiment, the present invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising (a) classifying each of a plurality of samples or individuals on the basis of one phenotypic or genotypic characteristic into a plurality of first classes; and (b) identifying within each of said first classes a set of informative genes or markers, wherein said set of informative genes or markers within each said first classes is unique to said class.

5.3.2 IDENTIFICATION OF MARKER SETS INFORMATIVE FOR PATIENT SUBSETS

[0081] Once a patient subset is identified, markers, such as genes, informative for a particular outcome, such as prognosis, may be identified. In one embodiment, the method for identifying marker sets is as follows. This example describes the use of genes and gene-derived nucleic acids as markers; however, proteins or other cellular constituents may be used as markers of the condition.

[0082] After extraction and labeling of target polynucleotides, the expression of a plurality of markers, such as genes, in a sample X is compared to the expression of the plurality of markers in a standard or control. In one embodiment, the standard or control comprises target markers, such as polynucleotide molecules, derived from one or more samples from a plurality of normal individuals, or a plurality of individuals not exposed to a particular condition. For example, the control, or normal, individuals may be persons without the particular disease or condition of interest (*e.g.*, individuals not afflicted with breast cancer, where breast cancer is the disease of interest), or may be an individual not exposed to a particular environmental condition. The standard or control may also comprise target polynucleotide molecules, derived from one or more samples derived from individuals having a different form or stage of the same disease; a different disease or different condition, or individuals exposed or subjected to a different condition, than the individual from which sample X was obtained. The control may be a sample, or set of samples, taken from the individual at an earlier time, for example, to assess the progression of a condition, or the response to a course of therapy.

[0083] In a preferred embodiment, the standard or control is a pool of target polynucleotide molecules. However, where protein levels, or the levels of any other relevant biomolecule, are to be compared, the pool may be a pool of proteins or the relevant biomolecule. In a preferred embodiment in the context of breast cancer, the pool comprises samples taken from a number of individuals having sporadic-type tumors.

[0084] In another preferred embodiment, the pool comprises an artificially-generated population of nucleic acids designed to approximate the level of nucleic acid derived from each marker found in a pool of marker-derived nucleic acids derived from tumor samples. In another embodiment, the pool, also called a "mathematical sample pool," is represented by a set of expression values, rather than a set of physical polynucleotides; the level of expression of relevant markers in a sample from an individual with a condition, such as a disease, is compared to values representing control levels of expression for the same markers in the mathematical sample pool. Such a control may be a set of values stored on a computer. Such artificial or mathematical controls may be constructed for any condition of interest.

[0085] In another embodiment specific to breast cancer, the pool is derived from normal or breast cancer cell lines or cell line samples. In a preferred embodiment, the pool comprises samples taken from individuals within a specific patient subset, *e.g.*, “ER+, ER/AGE high” individuals, wherein each of said individuals has a good prognosis, or each of said individuals has a poor prognosis. Of course, where, for example, expressed proteins are used as markers, the proteins are obtained from the individual’s sample, and the standard or control could be a pool of proteins from a number of normal individuals, or from a number of individuals having a particular state of a condition, such as a pool of samples from individuals having a particular prognosis of breast cancer.

[0086] The comparison may be accomplished by any means known in the art. For example, expression levels of various markers may be assessed by separation of target polynucleotide molecules (*e.g.*, RNA or cDNA) derived from the markers in agarose or polyacrylamide gels, followed by hybridization with marker-specific oligonucleotide probes. Alternatively, the comparison may be accomplished by the labeling of target polynucleotide molecules followed by separation on a sequencing gel. Polynucleotide samples are placed on the gel such that patient and control or standard polynucleotides are in adjacent lanes. Comparison of expression levels is accomplished visually or by means of densitometer. In a preferred embodiment, the expression of all markers is assessed simultaneously by hybridization to a microarray. In each approach, markers meeting certain criteria are identified as informative for the prognosis of breast cancer.

[0087] Marker genes are selected based upon significant difference of expression in a condition, such as a disease, as compared to a standard or control condition. Marker genes may be screened, for example, by determining whether they show significant variation within a set of samples of interest. Genes that do not show a significant amount of variation within the set of samples are presumed not to be informative for the disease or condition, and are not selected as markers for the disease or condition. Genes showing significant variation within the sample set are candidate informative genes for the disease or condition. The degree of variation may be estimated by calculating the difference of the expression of the gene, or ratio of expression between sample and control, within the set of samples. The expression, or ratio of expressions, may be transformed by any means, *e.g.*, linear or log transformation. Selection may be made based upon either significant up- or down regulation of the marker in the patient sample. Selection may also be made by calculation of the statistical significance (*i.e.*, the p-value) of the correlation between the expression of the marker and the disease and condition. Preferably, both selection criteria are used. Thus, in one embodiment of the

present invention, markers associated with prognosis of breast cancer within a patient subset are selected where the markers show both more than two-fold change (increase or decrease) in expression as compared to a standard, and the p-value for the correlation between the existence of breast cancer and the change in marker expression is no more than 0.01 (*i.e.*, is statistically significant).

[0088] In the context of the present invention, “good prognosis” indicates a desired outcome for a particular condition, especially a particular disease, and “poor prognosis” indicates an undesired outcome of the condition. For example, where the condition is cancer, a “good prognosis” may mean partial or complete remission, and “poor prognosis” may mean reappearance of the cancer after treatment. What constitutes “good prognosis” and “poor prognosis” is specific to the condition of interest, for example, specific to the particular cancer an individual suffers. For example, “good prognosis” for pancreatic cancer may be survival for one or two years after initial diagnosis, while “good prognosis” for Hodgkin’s disease may be survival for five years or more. In the specific example of breast cancer, “good prognosis” means the likelihood of non-reoccurrence of metastases within a period of 1, 2, 3, 4, 5 or more years after initial diagnosis, and “poor prognosis” means the likelihood of reoccurrence of metastasis within that period. In a more specific example, “good prognosis” means the likelihood of non-reoccurrence of metastases within 5 years after initial diagnosis, and “poor prognosis” means the likelihood of reoccurrence of metastasis within that period.

[0089] In a more specific embodiment for cancer, for example, breast cancer, using a number of breast cancer tumor samples, markers are identified by calculation of correlation coefficients ρ between the clinical category or clinical parameter(s) \vec{c} and the linear, logarithmic or any transform of the expression ratio \vec{r} across all samples for each individual gene. Specifically, the correlation coefficient may be calculated as:

$$[0090] \quad \rho = (\vec{c} \bullet \vec{r}) / (\|\vec{c}\| \cdot \|\vec{r}\|) \quad \text{Equation (1)}$$

[0091] Markers for which the coefficient of correlation exceeds a cutoff are identified as prognosis-informative markers specific for a particular clinical type, *e.g.*, good prognosis, within a given patient subset. Such a cutoff or threshold may correspond to a certain significance of discriminating genes obtained by Monte Carlo simulations. The threshold depends upon the number of samples used; the threshold can be calculated as $3 \times 1/\sqrt{n-3}$, where $1/\sqrt{n-3}$ is the distribution width and n = the number of samples. In a specific

embodiment, markers are chosen if the correlation coefficient is greater than about 0.3 or less than about -0.3.

[0092] Next, the significance of the correlation is calculated. This significance may be calculated by any statistical means by which such significance is calculated. In a specific example, a set of correlation data is generated using a Monte-Carlo technique to randomize the association between the expression difference of a particular marker and the clinical category. The frequency distribution of markers satisfying the criteria in the Monte-Carlo runs is used to determine whether the number of markers selected by correlation with clinical data is significant.

[0093] Once a marker set is identified, the markers may be rank-ordered in order of significance of discrimination. One means of rank ordering is by the amplitude of correlation between the change in gene expression of the marker and the specific condition being discriminated. Another, preferred, means is to use a statistical metric. In a specific embodiment, the metric is a t-test-like statistic:

$$[0094] \quad t = \frac{(\langle x_1 \rangle - \langle x_2 \rangle)}{\sqrt{[\sigma_1^2(n_1 - 1) + \sigma_2^2(n_2 - 1)] / (n_1 + n_2 - 1) / (1/n_1 + 1/n_2)}} \quad \text{Equation (2)}$$

[0095] In this equation, $\langle x_1 \rangle$ is the error-weighted average of the log ratio of transcript expression measurements within a first clinical group (e.g., good prognosis), $\langle x_2 \rangle$ is the error-weighted average of log ratio within a second, related clinical group (e.g., poor prognosis), σ_1 is the variance of the log ratio within the first clinical group (e.g., good prognosis), n_1 is the number of samples for which valid measurements of log ratios are available, σ_2 is the variance of log ratio within the second clinical group (e.g., poor prognosis), and n_2 is the number of samples for which valid measurements of log ratios are available. The t -value represents the variance-compensated difference between two means.

[0096] The rank-ordered marker set may be used to optimize the number of markers in the set used for discrimination. This is accomplished generally in a "leave one out" method as follows. In a first run, a subset, for example five, of the markers from the top of the ranked list is used to generate a template, where out of X samples, $X-1$ are used to generate the template, and the status of the remaining sample is predicted. This process is repeated for every sample until every one of the X samples is predicted once. In a second run, additional markers, for example five additional markers, are added, so that a template is now generated from 10 markers, and the outcome of the remaining sample is predicted. This process is

repeated until the entire set of markers is used to generate the template. For each of the runs, type 1 error (false negative) and type 2 errors (false positive) are counted; the optimal number of markers is that number where the type 1 error rate, or type 2 error rate, or preferably the total of type 1 and type 2 error rate is lowest.

[0097] For prognostic markers, validation of the marker set may be accomplished by an additional statistic, a survival model. This statistic generates the probability of tumor distant metastases as a function of time since initial diagnosis. A number of models may be used, including Weibull, normal, log-normal, log logistic, log-exponential, or log-Rayleigh (Chapter 12 “Life Testing”, S-PLUS 2000 GUIDE TO STATISTICS, Vol. 2, p. 368 (2000)). For the “normal” model, the probability of distant metastases P at time t is calculated as

$$[0098] \quad P = \alpha \times \exp(-t^2/\tau^2) \quad \text{Equation (3)}$$

[0099] where α is fixed and equal to 1, and τ is a parameter to be fitted and measures the “expected lifetime”.

[00100] It is preferable that the above marker identification process be iterated one or more times by excluding one or more samples from the marker selection or ranking (*i.e.*, from the calculation of correlation). Those samples being excluded are the ones that can not be predicted correctly from the previous iteration. Preferably, those samples excluded from marker selection in this iteration process are included in the classifier performance evaluation, to avoid overstating the performance.

[00101] It will be apparent to those skilled in the art that the above methods, in particular the statistical methods described above, are not limited to the identification of markers associated with the prognosis of breast cancer within a particular patient subset, but may be used to identify set of marker genes associated with any phenotype or condition, or with any subset of a phenotype or condition defined by one or more characteristics of the phenotype or condition. The phenotype or condition can be the presence or absence of a disease such as cancer, or the presence or absence of any identifying clinical condition associated with that cancer. In the disease context, the phenotype may be a prognosis such as a survival time, probability of distant metastases of a disease condition, or likelihood of a particular response to a therapeutic or prophylactic regimen. The phenotype need not be cancer, or a disease; the phenotype may be a nominal characteristic associated with a healthy individual.

[00102] Thus, the invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics,

comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; (b) identifying within each of said first classes a first set of genes or markers informative for said condition, wherein said first set of genes or markers within each of said first classes is unique to said class relative to other classes. In a specific embodiment, samples or individuals in at least one of said first classes are additionally classified on the basis of a phenotypic or genotypic characteristic different from that used to distinguish said first classes into a plurality of second classes, and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other classes. In another embodiment, the invention provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes; (b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used to distinguish said plurality of first classes; (c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition, wherein said set of genes or markers is unique to said class relative to other classes. The invention further provides a method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising: (a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics; (b) identifying at least two first condition classes differentiable by said first characteristic; (c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and (d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

5.3.3 CLASSIFIER GENESETS FOR FIVE PATIENT SUBSETS

[00103] The present invention provides sets of markers useful for the prognosis of breast cancer. The markers were identified according to the above methods in specific subsets of individuals with breast cancer. Generally, the marker sets were identified within a population of breast cancer patients that had been first stratified into five phenotypic categories based on criteria relevant to breast cancer prognosis, including estrogen receptor (ER) status, lymph

node status, type of mutation(s) (*i.e.*, *BRCA1*-type or sporadic) and age at diagnosis. More specifically, patients, and tumors from which samples were taken, were classified as ER^- , sporadic (*i.e.*, being both estrogen receptor negative and having a non-*BRCA1*-type tumor); ER^- , *BRCA1* (*i.e.*, being both estrogen receptor negative and having a *BRCA1*-type tumor); ER^+ , *ER/AGE* high (*i.e.*, estrogen receptor positive with a high ratio of the log (ratio) of estrogen receptor gene expression to age); ER^+ , *ER/AGE* low, LN^+ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node positive); and ER^+ , *ER/AGE* low, LN^- (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node negative). The rationale for subdivision of the original patient set into these five subsets is detailed in the Examples (Section 6). The marker sets useful for each of the subsets above are provided in Tables 1-5, respectively.

Table 1: Geneset of 20 markers used to classify ER^- , sporadic individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correlation	Description	Sp_xref_keyword_list	SEQ ID
AF055033	IGFBP5	-2.12	0.88	0.54	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure	11
NM_000599	IGFBP5	-3.41	0.43	0.53	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure	51
L27560	IGFBP5	-4.55	0	0.52	EST	Hypothetical protein	29
AF052162	FLJ12443	-0.27	1.6	0.52	EST	Hypothetical protein	9
NM_001456	FLNA	-0.61	2.47	0.52	filamin A, alpha (actin binding protein 280)	Hypothetical protein, Actin-binding, Phosphorylation, Repeat, Polymorphism, Disease mutation	73
NM_002205	ITGA5	-0.37	2.08	0.49	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	Integrin, Cell adhesion, Receptor, Glycoprotein, Transmembrane, Signal, Calcium, Repeat	93
NM_013261	PPARGC1	0.09	1.54	0.47	peroxisome proliferative activated receptor, gamma, coactivator 1		231
NM_001605	AARS	0.39	2.36	0.51	alanyl-tRNA synthetase	Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding	77
X87949	HSPA5	-0.03	2.03	0.49	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	ATP-binding, Hypothetical protein, Endoplasmic reticulum, Signal	273
Contig50950_RC	NGEF	-1.17	3.2	0.52	neuronal guanine nucleotide exchange factor		337

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correlation	Description	Sp_xref_keyword_list	SEQ ID
NM_005689	ABCB6	-0.51	2.26	0.48	ATP-binding cassette, sub-family B (MDR/TAP), member 6	ATP-binding, Transport, Transmembrane, Mitochondrion, Inner membrane, Transit peptide, Hypothetical protein	187
NM_004577	PSPH	-0.56	3.05	0.51	phosphoserine phosphatase	Hydrolase, Serine biosynthesis, Magnesium, Phosphorylation	151
NM_003832	PSPHL	-2.08	2.18	0.5	phosphoserine phosphatase-like		131
NM_002422	MMP3	-0.96	2.54	0.5	matrix metalloproteinase 3 (stromelysin 1, progelatinase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Signal, Polymorphism, 3D-structure	101
Contig37562_RC		-3.42	-6.02	-0.59	ESTs		293
NM_018465	MDS030	-0.82	-3.28	-0.58	uncharacterized hematopoietic stem/progenitor cells protein MDS030	Hypothetical protein	267
Contig54661_RC		-0.79	-2.08	-0.54	ESTs		349
AB032969	KIAA1143	-0.6	-2.85	-0.53	KIAA1143 protein	Hypothetical protein	1
Contig55353_RC	KIAA1915	-0.27	-1.82	-0.47	KIAA1915 protein	Hypothetical protein	353
NM_005213	CSTA	2.11	-3.4	-0.49	cystatin A (stefin A)	Thiol protease inhibitor, 3D-structure	175

Table 2. Geneset of 10 markers used to classify ER⁻, *BRCA1* individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correlation	Sequence name	Description	Sp_xref_keyword_list	SEQ ID
AF005487		6.08	0.5	-0.79	HLA-DRB6	Homo sapiens MHC class II antigen (DRB6) mRNA, HLA-DRB6*0201 allele, sequence.	MHC	3
Contig50728_RC		4.02	0.25	-0.77		ESTs, Weakly similar to S26650 DNA-binding protein 5 - human [H.sapiens]		333
Contig53598_RC		8.41	3.26	-0.77	FLJ11413	hypothetical protein FLJ11413	Hypothetical protein	343
NM_002888	RARR ES1	6.9	0.05	-0.87	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	Receptor, Transmembrane, Signal-anchor	109

NM_005218	DEFB1	5.14	-3.02	-0.81	DEFB1	defensin, beta 1	Antibiotic, Signal, 3D-structure	177
U17077	BENE	2.72	-1.72	-0.77	BENE	BENE protein	Transmembrane	271
Contig14683_RC		1.29	-2.31	-0.74		ESTs		279
Contig53641_RC		-3.29	4.23	0.75	MAGE-E1	MAGE-E1 protein	Hypothetical protein	345
Contig56678_RC		-6.7	-9.73	-0.82		ESTs, Highly similar to THYA_HUMAN Prothymosin alpha [H.sapiens]		357
NM_005461	KRML	0.88	-3.38	-0.75	MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	Transcription regulation, Repressor, DNA-binding, Nuclear protein, Hypothetical protein	181

Table 3. Geneset of 50 markers used to classify ER+, ER/AGE high individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_003600	STK15	-2.93	2.08	0.8	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase	125
NM_003158	STK6	-1.57	1.42	0.78	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase	113
NM_007019	UBCH10	-2.98	2.62	0.81	ubiquitin-conjugating enzyme E2C	Hypothetical protein, Ubl conjugation pathway, Ligase, Multigene family, Mitosis, Cell cycle, Cell division	217
NM_013277	ID-GAP	-2.43	2.43	0.77	Rac GTPase activating protein 1	Hypothetical protein	233
NM_004336	BUB1	-2.04	1.39	0.77	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	Transferase, Serine/threonine-protein kinase, ATP-binding, Cell cycle, Nuclear protein, Mitosis, Phosphorylation, Polymorphism	147

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_006607	PTTG2	-1.71	1.49	0.72	pituitary tumor- transforming 2		211
AK001166	FLJ11252	-1.33	0.99	0.71	hypothetical protein FLJ11252	Hypothetical protein	13
NM_004701	CCNB2	-4.62	2.01	0.81	cyclin B2	Cyclin, Cell cycle, Cell division, Mitosis	153
Contig57584_RC		-3.68	2.04	0.78	likely ortholog of mouse gene rich cluster, C8 gene		359
NM_006845	KNSL6	-4.13	1.05	0.73	kinesin-like 6 (mitotic centromere- associated kinesin)	Hypothetical protein, Motor protein, Microtubules, ATP- binding, Coiled coil, Nuclear protein	215
Contig38901_RC		-3.08	1.15	0.75	hypothetical protein MGC45866	Hypothetical protein	299
NM_018410	DKFZp76 2E1312	-4.38	1.49	0.75	hypothetical protein DKFZp762E1312	Hypothetical protein	263
NM_003981	PRC1	-3.52	2.17	0.78	protein regulator of cytokinesis 1		133
NM_001809	CENPA	-5.04	0.98	0.75	centromere protein A, 17kDa	Hypothetical protein, Chromosomal protein, Nuclear protein, DNA- binding, Centromere, Antigen	81
NM_003504	CDC45L	-2.67	1.22	0.73	CDC45 cell division cycle 45-like (S. cerevisiae)	DNA replication, Cell cycle, Nuclear protein, Cell division	123
Contig41413_RC		-5.43	2.15	0.74	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron	305
NM_004217	STK12	-2.17	0.73	0.72	serine/threonine kinase 12	Hypothetical protein, ATP- binding, Kinase, Serine/threonine- protein kinase, Transferase	143
NM_002358	MAD2L1	-2.65	2.27	0.83	MAD2 mitotic arrest deficient-like 1 (yeast)	Cell cycle, Mitosis, Nuclear protein, 3D-structure	99
NM_014321	ORC6L	-2.73	1.8	0.75	origin recognition complex, subunit 6 homolog-like (yeast)	Hypothetical protein, DNA replication, Nuclear protein, DNA- binding	241
NM_012291	KIAA0165	-1.52	1.55	0.71	extra spindle poles like 1 (S. cerevisiae)	Hypothetical protein	229

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_004203	PKMYT1	-3.64	2.2	0.7	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase, Transcription regulation, DNA- binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene	137
M96577	E2F1	-2.14	1.42	0.75	E2F transcription factor 1	Transcription regulation, Activator, DNA- binding, Nuclear protein, Phosphorylation, Cell cycle, Apoptosis, Polymorphism	33
NM_002266	KPNA2	-3.77	1.78	0.71	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism	95
Contig31288_RC		-2.63	0.7	0.68	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]		289
NM_014501	E2-EPF	-1.55	1.93	0.7	ubiquitin carrier protein	Ubl conjugation pathway, Ligase, Multigene family	247
NM_001168	BIRC5	-5.76	2.01	0.78	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor	63
NM_003258	TK1	-4.57	1.38	0.71	thymidine kinase 1, soluble	Transferase, Kinase, DNA synthesis, ATP- binding	115
NM_001254	CDC6	-2.46	0.28	0.72	CDC6 cell division cycle 6 homolog (S. cerevisiae)	ATP-binding, Cell division	67
NM_004900	DJ742C19 .2	-2.96	0.13	0.69	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Hydrolase	161
NM_004702	CCNE2	-3.12	2.13	0.81	cyclin E2	Cell cycle, Cell division, Cyclin, Hypothetical protein, Phosphorylation, Alternative splicing, Nuclear protein	155

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
AL160131		-3.07	2.42	0.7	hypothetical protein MGC861	Hypothetical protein	21
NM_016359	LOC5120 3	-3.22	2.61	0.76	nucleolar protein ANKT	Hypothetical protein, Nuclear protein	253
NM_004856	KNSL5	-1.52	1.1	0.71	kinesin-like 5 (mitotic kinesin-like protein 1)	Motor protein, Cell division, Microtubules, ATP- binding, Coiled coil, Mitosis, Cell cycle, Nuclear protein	159
NM_000057	BLM	-1.54	0.76	0.71	Bloom syndrome	Hydrolase, Helicase, ATP- binding, DNA- binding, Nuclear protein, DNA replication, Disease mutation	35
NM_018455	BM039	-2.44	1.18	0.7	uncharacterized bone marrow protein BM039		265
NM_002106	H2AFZ	-2.49	1.53	0.72	H2A histone family, member Z	Chromosomal protein, Nucleosome core, Nuclear protein, DNA-binding, Multigene family	91
Contig64688		-2.68	3.1	0.73	hypothetical protein FLJ23468	Hypothetical protein	365
Contig44289_RC		-1.65	1.6	0.67	ESTs		315
Contig28552_RC		-1.37	1.53	0.68	diaphanous homolog 3 (Drosophila)	Hypothetical protein, Coiled coil, Repeat, Alternative splicing	281
Contig46218_RC		-1.31	1.56	0.68	ESTs, Weakly similar to T19201 hypothetical protein C11G6.3 - Caenorhabditis elegans [C. elegans]		321
Contig28947_RC		-1.3	0.98	0.67	cell division cycle 25A	Hypothetical protein, Cell division, Mitosis, Hydrolase, Alternative splicing, Multigene family, 3D-structure	283
NM_016095	LOC5165 9	-1.4	2.13	0.67	HSPC037 protein	Hypothetical protein	249
NM_003090	SNRPA1	-3.26	0.95	0.7	small nuclear ribonucleoprotein polypeptide A'	Hypothetical protein, Nuclear protein, RNA- binding, Ribonucleoprotein, Leucine-rich repeat, Repeat, 3D-structure	111

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword _list	SEQ ID
NM_002811	PSMD7	-2.48	1.89	0.7	proteasome (prosome, macropain) 26S subunit, non-ATPase, 7 (Mov34 homolog)	Proteasome	107
Contig38288_RC		-2.34	0.97	0.67	hypothetical protein DKFZp762A2013	Hypothetical protein	297
NM_003406	YWHAZ	-1.5	2.79	0.68	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	Brain, Neurone, Phosphorylation, Acetylation, Multigene family, 3D-structure	121
AL137540	NTN4	2.13	-4.61	-0.69	netrin 4	Hypothetical protein, Laminin EGF-like domain, Signal	19
AL049367		1.9	-3.2	-0.68	EST	Transducer, Prenylation, Lipoprotein, Multigene family, Acetylation	15
NM_013409	FST	1.04	-5.78	-0.69	folistatin	Glycoprotein, Repeat, Signal, Alternative splicing	235
NM_000060	BTD	3.1	-1.45	-0.67	biotinidase	Hydrolase, Glycoprotein, Signal, Disease mutation	37

Table 4. Geneset of 50 markers used to classify ER+, ER/AGE low, LN+ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
--------------------------	------	---------------------	---------------------	------------------	-------------	--------------------------	--------

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_006417	MTAP44	-1.5	3	0.69	Fc fragment of IgG, low affinity IIb, receptor for (CD32)	Hydrolase, Hypothetical protein, Immunoglobulin domain, IgG- binding protein, Receptor, Transmembrane, Glycoprotein, Signal, Repeat, Multigene family, Polymorphism, NAD, One-carbon metabolism, Serine protease, Zymogen, Protease, Alternative splicing, Chromosomal translocation, Proto-oncogene, Galactin, Lectin, Antigen	205
NM_006820	GS3686	-4.3	4.06	0.69	chromosome 1 open reading frame 29	Hypothetical protein	213
NM_001548	IFIT1	-3.4	4.27	0.71	Interferon-induced protein with tetrapeptide repeats 1	Repeat, TPR repeat, Interferon induction	75
Contig41538_RC		-2.5	3.16	0.68	ESTs, Moderately similar to hypothetical protein FLJ20489 [<i>Homo sapiens</i>]		307
NM_016816	OAS1	-1.7	3.29	0.75	2',5'-oligoadenylate synthetase 1, 40/46kDa	RNA-binding, Transferase, Nucleotidyltransfer ase, Interferon induction, Alternative splicing	255
Contig51660_RC		-2.1	2.65	0.66	28kD interferon responsive protein	Transmembrane	339
Contig43645_RC		-4.8	1.44	0.63	<i>Homo sapiens</i> , clone IMAGE:4428577, mRNA, partial cds	Hypothetical protein	313
AF026941		-4.6	2.71	0.63	EST, Weakly similar to 2004399A chromosomal protein [<i>Homo sapiens</i>]	Hypothetical protein	5
NM_007315	STAT1	-3.5	1.8	0.59	signal transducer and activator of transcription 1, 91kDa	Transcription regulation, DNA- binding, Nuclear protein, Phosphorylation, SH2 domain, Alternative splicing, 3D-structure	225

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_002038	G1P3	-4.1	5.64	0.79	interferon, alpha-inducible protein (clone IFI-6-16)	Interferon induction, Transmembrane, Signal, Alternative splicing	85
NM_005101	ISG15	-5.6	5.34	0.77	interferon-stimulated protein, 15 kDa	Interferon induction, Repeat	169
NM_002462	MX1	-6.1	0.83	0.56	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	Hypothetical protein, Interferon induction, GTP-binding, Multigene family, Antiviral	103
NM_005532	IFI27	-5.8	2.81	0.59	interferon, alpha-inducible protein 27	Interferon induction, Transmembrane	183
NM_002346	LY6E	-2.1	3.58	0.75	lymphocyte antigen 6 complex, locus E	Signal, Antigen, Multigene family, Membrane, GPI-anchor	97
NM_016817	OAS2	-3.6	1.89	0.59	2'-5'-oligoadenylate synthetase 2, 69/71kDa	RNA-binding, Transferase, Nucleotidyltransferase, Repeat, Interferon induction, Alternative splicing, Myristate	257
Contig44909_RC		-2.3	1.13	0.55	hypothetical protein BC012330	Hypothetical protein	317
NM_017414	USP18	-4.1	3.37	0.72	ubiquitin specific protease 18	Ubl conjugation pathway, Hydrolase, Thiol protease, Multigene family	259
NM_004029	IRF7	-2.4	3.67	0.66	interferon regulatory factor 7	Collagen, Transcription regulation, DNA-binding, Nuclear protein, Activator, Alternative splicing	135
NM_004335	BST2	-3.2	3.22	0.57	bone marrow stromal cell antigen 2	Transmembrane, Glycoprotein, Signal-anchor, Polymorphism	145
NM_002759	PRKR	-2.4	1.8	0.58	protein kinase, interferon-inducible double stranded RNA dependent	Transferase, Serine/threonine-protein kinase, ATP-binding, Repeat, Phosphorylation, Interferon induction, RNA-binding, 3D-structure	105

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_006332	IFI30	-3.8	2.65	0.64	interferon, gamma-inducible protein 30	Oxidoreductase, Interferon induction, Glycoprotein, Lysosome, Signal, Hypothetical protein	203
NM_009587	LGALS9	-3.2	2.08	0.6	lectin, galactoside-binding, soluble, 9 (galectin 9)	Galaptin, Lectin, Repeat, Alternative splicing	227
NM_003641	IFITM1	-2.4	5.54	0.63	interferon induced transmembrane protein 1 (9-27)	Interferon induction, Transmembrane	127
NM_017523	HSXIAPA F1	-1	2.84	0.7	XIAP associated factor-1	Hypothetical protein	261
NM_014314	RIG-I	-1.3	3.55	0.62	RNA helicase	ATP-binding, Helicase, Hydrolase, Hypothetical protein	239
Contig47563_RC		-2.2	3.11	0.56	ESTs		325
AI497657_RC		-4.4	5.61	0.74	guanine nucleotide binding protein 4	Transducer, Prenylation, Lipoprotein, Multigene family	335
NM_000735	CGA	-4.3	2.5	0.58	glycoprotein hormones, alpha polypeptide	Hormone, Glycoprotein, Signal, 3D-structure	53
NM_004988	MAGEA1	-1.4	6.31	0.64	melanoma antigen, family A, 1 (directs expression of antigen MZ2-E)	Antigen, Multigene family, Polymorphism, Tumor antigen	163
Contig54242_RC		-1.2	4.1	0.65	chromosome 17 open reading frame 26	Hypothetical protein	347
NM_004710	SYNGR2	-1.4	3.01	0.54	synaptogyrin 2	Transmembrane	157
NM_001168	BIRC5	-3.7	3.39	0.64	baculoviral IAP repeat-containing 5 (survivin)	Hypothetical protein, Protease, Receptor, Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure	63
Contig41413_RC		-4.4	2.61	0.57	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron	305

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
NM_004203	PKMYT1	-3.4	3.79	0.6	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene	137
Contig48913_RC		-3.1	1.72	0.55	<i>Homo sapiens</i> , Similar to hypothetical protein PRO1722, clone MGC:15692 IMAGE:3351479, mRNA, complete cds		327
NM_005804	DDXL	-2.5	1.42	0.58	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 39	ATP-binding, Helicase, Hydrolase, Hypothetical protein	191
NM_016359	LOC51203	-1.7	3.6	0.57	nucleolar protein ANKT	Hypothetical protein, Nuclear protein	253
NM_001645	APOC1	-2.9	3.43	0.58	apolipoprotein C-I	Plasma, Lipid transport, VLDL, Signal, 3D-structure, Polymorphism	79
Contig37895_RC		-2	2.05	0.55	ESTs		295
NM_005749	TOB1	-1.3	4.96	0.59	transducer of ERBB2, 1	Phosphorylation	189
NM_000269	NME1	-1.3	2.98	0.55	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti-oncogene, Disease mutation	39
NM_014462	LSM1	-1	4.5	0.57	Lsm1 protein	Nuclear protein, Ribonucleoprotein, mRNA splicing, mRNA processing, RNA-binding	245
Contig31221_RC		-1.4	3.83	0.56	HTPAP protein		287
NM_005326	HAGH	-1.9	4.29	0.57	hydroxyacyl glutathione hydrolase	Hydrolase, Zinc, 3D-structure	179
Contig42342_RC		0.78	-3.2	-0.6	<i>Homo sapiens</i> cDNA FLJ39417 fis, clone PLACE6016942	Hypothetical protein	311
AL137540	NTN4	2.24	-3.9	-0.6	netrin 4	Laminin EGF-like domain, Signal, Hypothetical protein	19
Contig40434_RC		1.64	-5.6	-0.6	wingless-type MMTV integration site family, member 5A	Developmental protein, Glycoprotein, Signal	301

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword _list	SEQ ID
Contig1632_RC		1.03	-3.9	-0.6	hypothetical protein MGC17921	Hypothetical protein	275
NM_014246	CELSR1	0.95	-4.6	-0.6	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, <i>Drosophila</i>)	G-protein coupled receptor, Transmembrane, Glycoprotein, EGF- like domain, Calcium-binding, Laminin EGF-like domain, Repeat, Developmental protein, Hydroxylation, Signal, Alternative splicing, Hypothetical protein	237
NM_005139	ANXA3	1.26	-6.2	-0.6	annexin A3	Annexin, Calcium/phospholi pid-binding, Repeat, Phospholipase A2 inhibitor, 3D- structure, Polymorphism	171

Table 5. Geneset of 65 markers used to classify ER⁺, ER/AGE low, LN⁻ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
M55914	MPB1	-2.82	1.25	0.5	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein	31
NM_005945	MPB1	-3.06	1.19	0.49	ENO1	Homo sapiens enolase 1, (alpha) (ENO1), mRNA.	Glycolysis, Hypothetical protein, Lyase, Magnesium, DNA-binding, Transcription regulation, Repressor, Nuclear protein, Multigene family	193

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_001428	ENO1	-2.53	1.18	0.46	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein	71
NM_001216	CA9	-4.72	1.49	0.6	CA9	carbonic anhydrase IX	Lyase, Zinc, Transmembrane , Glycoprotein, Antigen, Signal, Nuclear protein, Polymorphism	65
NM_001124	ADM	-5.68	2.99	0.56	ADM	Adrenomedullin	Hormone, Amidation, Cleavage on pair of basic residues, Signal	61
NM_000584	IL8	-2.45	2.04	0.54	IL8	interleukin 8	Cytokine, Chemotaxis, Inflammatory response, Signal, Alternative splicing, 3D- structure	49
D25328	PFKP	-4.19	3.29	0.56	PFKP	Phosphofructo- kinase, platelet	Kinase, Transferase, Glycolysis, Repeat, Allosteric enzyme, Phosphorylation, Magnesium, Multigene family	25
NM_006096	NDRG1	-5.45	5.97	0.77	NDRG1	N-myc downstream regulated gene 1	Hypothetical protein, Nuclear protein, Repeat	199
NM_004994	MMP9	-5.53	1.07	0.49	MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Repeat, Signal, Polymorphism, 3D-structure	165

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_003311	TSSC3	-4.57	5.58	0.68	TSSC3	tumor suppressing subtransferable candidate 3		117
NM_006086	TUBB4	-5.19	2.85	0.59	TUBB4	tubulin, beta, 4	G-protein coupled receptor, Transmembrane, Glycoprotein, Phosphorylation, Lipoprotein, Palmitate, Polymorphism, Hypothetical protein, GTP-binding, Receptor, Microtubules, Multigene family	197
NM_006115	PRAME	-4.48	2.77	0.61	PRAME	preferentially expressed antigen in melanoma	Antigen	201
NM_004345	CAMP	-2.02	1.37	0.49	CAMP	cathelicidin antimicrobial peptide	Antibiotic, Signal	149
NM_018455	BM039	-2.34	0.76	0.47	BM039	uncharacterized bone marrow protein BM039		265
Contig49169_RC		-1.17	1.5	0.46	SV39H2	suppressor of variegation 3-9 (Drosophila) homolog 2; hypothetical protein FLJ23414	Hypothetical protein, Nuclear protein	329
Contig45032_RC		-1.37	0.77	0.45	FLJ14813	hypothetical protein FLJ14813	Hypothetical protein, ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase	319
NM_000917	P4HA1	-1.54	4.31	0.62	P4HA1	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I	Dioxygenase, Collagen, Oxidoreductase, Iron, Vitamin C, Alternative splicing, Glycoprotein, Endoplasmic reticulum, Signal	57
NM_002046	GAPD	-2.51	3.42	0.6	GAPD	glyceraldehyde-3-phosphate dehydrogenase	Glycolysis, NAD, Oxidoreductase, Hypothetical protein, Multigene family	87

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_000365	TPI1	-1.81	2.94	0.56	TPI1	triosephosphate isomerase 1	Fatty acid biosynthesis, Gluconeogenesis, Glycolysis, Isomerase, Pentose shunt, Disease mutation, Polymorphism, 3D-structure	45
NM_014364	GAPDS	-1.08	2.88	0.58	GAPDS	glyceraldehyde-3- phosphate dehydrogenase, testis-specific	Glycolysis, Oxidoreductase, NAD	243
NM_005566	LDHA	-2.01	4.01	0.59	LDHA	lactate dehydrogenase A	Oxidoreductase, NAD, Glycolysis, Multigene family, Disease mutation, Polymorphism	185
NM_000291	PGK1	-2.28	1.68	0.51	PGK1	phosphoglycerate kinase 1	Kinase, Transferase, Multigene family, Glycolysis, Acetylation, Disease mutation, Polymorphism, Hereditary hemolytic anemia	41
NM_016185	LOC511 55	-2.33	2.82	0.59	HN1	hematological and neurological expressed 1		251
NM_001168	BIRC5	-4.33	2.78	0.55	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D- structure, Hypothetical protein, Protease, Receptor	63
NM_002266	KPNA2	-3.75	1.34	0.47	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism	95
Contig31288_RC		-2.1	1.27	0.5		ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]		289

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_000269	NME1	-2.15	3.43	0.55	NME1	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP- binding, Nuclear protein, Anti- oncogene, Disease mutation	39
NM_003158	STK6	-1.23	1.73	0.45	STK6	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine -protein kinase, Transferase	113
NM_007274	HBACH	-1.83	2.73	0.51	BACH	brain acyl-CoA hydrolase	Hydrolase, Serine esterase, Repeat	223
Contig55188_RC		-2.36	3.28	0.47	FLJ22341	hypothetical protein FLJ22341	Hypothetical protein	351
NM_002061	GCLM	-1.06	1.76	0.48	GCLM	glutamate-cysteine ligase, modifier subunit	Ligase, Glutathione biosynthesis	89
NM_004207	SLC16A 3	-3.11	5.07	0.67	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Transport, Symport, Transmembrane , Multigene family	139
NM_000582	SPP1	-5.09	5.47	0.53	SPP1	secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	Hypothetical protein, Glycoprotein, Sialic acid, Biom mineralizatio n, Cell adhesion, Phosphorylation, Signal, Alternative splicing	47
NM_001109	ADAM8	-2.5	3.74	0.45	ADAM8	a disintegrin and metalloproteinase domain 8	Hydrolase, Metalloprotease, Zinc, Signal, Glycoprotein, Transmembrane , Antigen	59
D50402	SLC11A 1	-1.05	3.46	0.53	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	Transport, Iron transport, Transmembrane , Glycoprotein, Macrophage, Polymorphism	27
AL080235	DKFZP5 86E162 1	-1.23	1.96	0.51	RIS1	Ras-induced senescence 1	Hypothetical protein	17
Contig40552_RC		-1.26	3.96	0.54	FLJ25348	hypothetical protein FLJ25348	Hypothetical protein	303
Contig52490_RC		-0.64	3.33	0.61	LOC11623 8	hypothetical protein BC014072		341
NM_006461	DEEPE ST	-2.1	1.85	0.46	SPAG5	sperm associated antigen 5	Hypothetical protein	207

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
Contig56503_RC		-4.3	3.39	0.55	MGC9753	hypothetical gene MGC9753	Hypothetical protein	355
Contig63525		-1.91	3.34	0.5	FLJ13352	hypothetical protein FLJ13352	Hypothetical protein	363
NM_001909	CTSD	-0.83	4.6	0.51	CTSD	cathepsin D (lysosomal aspartyl protease)	Hydrolase, Aspartyl protease, Glycoprotein, Lysosome, Signal, Zymogen, Polymorphism, Alzheimer's disease, 3D- structure	83
NM_005063	SCD	-2.57	5.15	0.48	SCD	stearoyl-CoA desaturase (delta- 9-desaturase)	Hypothetical protein, Endoplasmic reticulum, Fatty acid biosynthesis, Iron, Oxidoreductase, Transmembrane	167
NM_005165	ALDOC	-2.43	5.02	0.48	ALDOC	aldolase C, fructose- biphosphate	Lyase, Schiff base, Glycolysis, Multigene family	173
NM_000363	TNNI3	-0.54	3.58	0.48	TNNI3	troponin I, cardiac	Hypothetical protein, Muscle protein, Actin- binding, Acetylation, Disease mutation, Cardiomyopathy , Receptor, Signal	43
AF035284		-1.63	3.28	0.47	FADS1	EST	Heme, Hypothetical protein	7
Contig30875_RC		-0.88	3	0.6		ESTs		285
NM_018487	HCA112	-0.7	3.54	0.58	HCA112	hepatocellular carcinoma- associated antigen 112	Hypothetical protein	269
NM_001323	CST6	-1.63	3.84	0.57	CST6	cystatin E/M	Thiol protease inhibitor, Signal, Glycoprotein	69

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
NM_006516	SLC2A1	-1.66	2.22	0.46	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	Transmembrane, Sugar transport, Transport, Glycoprotein, Multigene family, Disease mutation	209
NM_007267	LAK-4P	-1.04	3.28	0.61	EVIN1	expressed in activated T/LAK lymphocytes	Hypothetical protein	221
NM_004710	SYNGR2	-0.84	4.81	0.56	SYNGR2	synaptogyrin 2	Transmembrane	157
Contig63649_RC		-1.34	6.3	0.75		ESTs, Weakly similar to 2004399A chromosomal protein [Homo sapiens] [H.sapiens]		361
NM_003376	VEGF	-2.12	2.42	0.46	VEGF	vascular endothelial growth factor	Hypothetical protein, Mitogen, Angiogenesis, Growth factor, Glycoprotein, Signal, Heparin-binding, Alternative splicing, Multigene family, 3D-structure	119
NM_000799	EPO	-0.75	4.01	0.69	EPO	erythropoietin	Erythrocyte maturation, Glycoprotein, Hormone, Signal, Pharmaceutical, 3D-structure	55
NM_006014	DXS9879E	-1.85	3.44	0.54	DXS9879E	DNA segment on chromosome X (unique) 9879 expressed sequence		195
NM_007183	PKP3	-0.91	4.14	0.48	PKP3	plakophilin 3	Cell adhesion, Cytoskeleton, Structural protein, Nuclear protein, Repeat	219
D13642	SF3B3	-0.65	2.28	0.48	SF3B3	splicing factor 3b, subunit 3, 130kDa	Hypothetical protein, Spliceosome, mRNA processing, mRNA splicing, Nuclear protein	23
NM_003756	EIF3S3	-1.85	2.19	0.46	EIF3S3	eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa	Initiation factor, Protein biosynthesis	129

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keywo rd_list	SEQ ID
Contig47096_RC		-0.41	4.52	0.54	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	Kinase, Multifunctional enzyme, Transferase, Hydrolase, ATP-binding, Phosphorylation, Multigene family	323
NM_004209	SYNGR3	-0.31	3.67	0.53	SYNGR3	synaptogyrin 3	Transmembrane	141
Contig3464_RC		0.99	-5.81	-0.52		ESTs		277
Contig31646_RC		1.1	-7.76	-0.5	COL14A1	collagen, type XIV, alpha 1 (undulin)	Extracellular matrix, Glycoprotein, Hypothetical protein, Collagen, Signal	291
Contig49388_RC		1.73	-1.75	-0.51	FLJ13322	hypothetical protein FLJ13322	Hypothetical protein	331
Contig41887_RC		0.37	-5.74	-0.47	LOC124220	similar to common salivary protein 1	Hypothetical protein	309

[00104]

5.4 DIAGNOSTIC AND PROGNOSTIC METHODS

5.4.1 SAMPLE COLLECTION

[00105] In the present invention, markers, such as target polynucleotide molecules or proteins, are extracted from a sample taken from an individual afflicted with a condition such as breast cancer. The sample may be collected in any clinically acceptable manner, but must be collected such that marker-derived polynucleotides (*i.e.*, RNA) are preserved (if gene expression is to be measured) or proteins are preserved (if encoded proteins are to be measured). For example, mRNA or nucleic acids derived therefrom (*i.e.*, cDNA or amplified DNA) are preferably labeled distinguishably from standard or control polynucleotide molecules, and both are simultaneously or independently hybridized to a microarray comprising some or all of the markers or marker sets or subsets described above.

Alternatively, mRNA or nucleic acids derived therefrom may be labeled with the same label as the standard or control polynucleotide molecules, wherein the intensity of hybridization of each at a particular probe is compared. A sample may comprise any clinically relevant tissue sample, such as a tumor biopsy or fine needle aspirate, or a sample of bodily fluid, such as blood, plasma, serum, lymph, ascitic fluid, cystic fluid, urine or nipple exudate. The sample may be taken from a human, or, in a veterinary context, from non-human animals such as ruminants, horses, swine or sheep, or from domestic companion animals such as felines and canines.

[00106] Methods for preparing total and poly(A)+ RNA are well known and are described generally in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)).

[00107] RNA may be isolated from eukaryotic cells by procedures that involve lysis of the cells and denaturation of the proteins contained therein. Cells of interest include wild-type cells (*i.e.*, non-cancerous), drug-exposed wild-type cells, tumor- or tumor-derived cells, modified cells, normal or tumor cell line cells, and drug-exposed modified cells. Preferably, the cells are breast cancer tumor cells.

[00108] Additional steps may be employed to remove DNA. Cell lysis may be accomplished with a nonionic detergent, followed by microcentrifugation to remove the nuclei and hence the bulk of the cellular DNA. In one embodiment, RNA is extracted from cells of the various types of interest using guanidinium thiocyanate lysis followed by CsCl centrifugation to separate the RNA from DNA (Chirgwin *et al.*, *Biochemistry* 18:5294-5299 (1979)). Poly(A)+ RNA is selected by selection with oligo-dT cellulose (*see* Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)). Alternatively, separation of RNA from DNA can be accomplished by organic extraction, for example, with hot phenol or phenol/chloroform/isoamyl alcohol.

[00109] If desired, RNase inhibitors may be added to the lysis buffer. Likewise, for certain cell types, it may be desirable to add a protein denaturation/digestion step to the protocol.

[00110] For many applications, it is desirable to preferentially enrich mRNA with respect to other cellular RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA). Most mRNAs contain a poly(A) tail at their 3' end. This allows them to be enriched by affinity chromatography, for example, using oligo(dT) or poly(U) coupled to a solid support, such as cellulose or Sephadex™ (*see* Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)). Once bound, poly(A)+ mRNA is eluted from the affinity column using 2 mM EDTA/0.1% SDS.

[00111] The sample of RNA can comprise a plurality of different mRNA molecules, each different mRNA molecule having a different nucleotide sequence. In a specific embodiment, the mRNA molecules in the RNA sample comprise at least 5, 10, 15, 20, 25, 30, 40 or 50 different nucleotide sequences. More preferably, the mRNA molecules of the RNA sample

comprise mRNA molecules corresponding to each of the marker genes. In another specific embodiment, the RNA sample is a mammalian RNA sample.

[00112] In a specific embodiment, total RNA or mRNA from cells are used in the methods of the invention. The source of the RNA can be cells of a plant or animal, human, mammal, primate, non-human animal, dog, cat, mouse, rat, bird, yeast, eukaryote, prokaryote, etc. In specific embodiments, the method of the invention is used with a sample containing total mRNA or total RNA from 1×10^6 cells or less. In another embodiment, proteins can be isolated from the foregoing sources, by methods known in the art, for use in expression analysis at the protein level.

[00113] Probes to the homologs of the marker sequences disclosed herein can be employed preferably when non-human nucleic acid is being assayed.

[00114] The methods of the invention may employ any molecule suitable as a marker. For example, sets of proteins informative for a particular condition, including a disease, may be determined. As for gene-based markers, levels of variations of different proteins in samples may be determined for phenotypic or genotypic subsets of the condition, and proteins showing significant variation in either level (abundance) or activity, or both, may be identified in order to create a set of proteins informative for one or more of these subsets. Such proteins may be identified, for example, by use of gel electrophoresis, such as one-dimensional polyacrylamide gel electrophoresis, two-dimensional polyacrylamide gel electrophoresis, nondenaturing polyacrylamide gel electrophoresis; isoelectric focusing gels, etc., by use of antibody arrays, etc. Of course, the particular template(s) used to classify the individual depends upon the type(s) of cellular constituents used as markers. For example, where nucleic acids (*e.g.*, genes or nucleic acids derived from expressed genes) are used as markers, the template comprises nucleic acids (or the level of expression or abundance thereof); where proteins are used as markers, the template comprises proteins, for example, the level or abundance of those proteins in a set of individuals; etc.

5.4.2 USE OF PROGNOSTIC GENESETS FOR BREAST CANCER

[00115] According to the present invention, once genesets informative for a plurality of subsets of a condition are identified, an individual is classified into one of these subsets and a prognosis is made based on the expression of the genes, or their encoded proteins, in the geneset for that subset in a breast cancer tumor sample taken from the individual.

[00116] For example, a particular hypothetical condition has three relevant phenotypic characteristics, A, B and C. In this example, based on these characteristics, genesets

informative for prognosis of four patient subsets A^+B^+ ; $A^+B^-C^+$; $A^+B^-C^-$; and A^- are identified by the method described above. Thus, an individual having the condition would first be classified according to phenotypes A-C into one of the four patient subsets. In one embodiment, therefore, the invention provides for the classification of an individual having a condition into one of a plurality of patient subsets, wherein a set of genes informative for prognosis for the subset has been identified. A sample is then taken from the individual, and the expression of the prognostically-informative genes in the sample is analyzed and compared to a control. In various embodiments, the control is the average expression of informative genes in a pool of samples taken from good prognosis individuals classifiable into that patient subset; the average expression of informative genes in a pool of samples taken from poor prognosis individuals classifiable into that patient subset; a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

[00117] In another embodiment, a sample is taken from the individual, and the levels of expression of the prognostically-informative genes in the sample is analyzed. In one embodiment, the expression level of each gene can be compared to the expression level of the corresponding gene in a control of reference sample to determine a differential expression level. The expression profile comprising expression levels or differential expression levels of the plurality of genes is then compared to a template profile. In various embodiments, the template profile is a good prognosis template comprising the average expression of informative genes in samples taken from good prognosis individuals classifiable into that patient subset; or a poor prognosis template comprising the average expression of informative genes in samples taken from poor prognosis individuals classifiable into that patient subset; or a good prognosis profile comprising a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

[00118] In a specific embodiment, the condition is breast cancer, and the phenotypic, genotypic and/or clinical classes are: ER^- , *BRCA1* individuals; ER^- , sporadic individuals; ER^+ , *ER/AGE* high individuals; ER^+ , *ER/AGE* low, *LN*⁺ individuals; and ER^+ , *ER/AGE* low, *LN*⁻ individuals. In this embodiment, an individual may be classified as ER^+ or ER^- . If the individual is ER^- , the individual is additionally classified as having a *BRCA1*-type or sporadic tumor. ER^- individuals are thus classified as ER^- , *BRCA1* or ER^- , sporadic. Alternatively, if the individual is classified as ER^+ , the individual is additionally classified as having a high or low ratio of the log (ratio) of the level of expression of the gene encoding the estrogen receptor to the individual's age. Individuals having a low ratio are additionally

classified as LN+ or LN-. ER+ individuals are thus classified as ER+, ER/AGE high; ER+, ER/AGE low, LN+, or ER+, ER/AGE low, LN-. Of course, the individual's ER status, tumor type, age and LN status may be identified in any order, as long as the individual is classified into one of these five subsets.

[00119] Thus, in one embodiment, the invention provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression. In a specific embodiment, said condition is breast cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In an more specific embodiment, said classifying said individual with a condition as having a good prognosis or a poor prognosis is carried out by comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis or poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in a good prognosis control or a poor prognosis control, respectively, more strongly than would be expected by chance. In a more specific embodiment of the method, said plurality of patient subsets comprises ER-, *BRCA1* individuals; ER-, sporadic individuals; ER+, ER/AGE high individuals; ER+, ER/AGE low, LN+ individuals; and ER+, ER/AGE low, LN- individuals. In another embodiment, said control is the average level of expression of each of said plurality of genes informative for prognosis in a pool of tumor samples from individuals classified into said subset who have a good prognosis or good outcome, or who have a poor prognosis or good outcome. In another specific embodiment, said control is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis.

[00120] It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining the level of expression in a sample from said individual of a plurality of the genes for which markers are listed in Tables 1-5; and comparing said expression to a control, where said control is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER⁺, ER/AGE high type; ER⁺, ER/AGE low, LN⁺ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said control is the average level of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a control level of expression of the same genes for each of Tables 1-5, and determining whether the level of expression of said genes in said sample is most similar to said control expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

[00121] In another embodiment, the invention provides a method of classifying an individual as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual as ER⁻, *BRCA1*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining the level of expression of a first plurality of genes in a cell sample taken from the individual relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻, wherein said individual is "ER/AGE high" if the ratio of ER expression to age exceeds a predetermined value, and "ER/AGE low" if the ratio of ER expression to age does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCA1*, and said

plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises (*i.e.*, contains at least) 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[00122] In one embodiment, the invention provides a method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition; (b) determining the levels of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual, optionally relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and (c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said levels of expression. In a specific embodiment, said condition is breast cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis. In a more specific embodiment, said classifying said individual with a condition as having a good prognosis or a poor prognosis is carried out by comparing the patient's expression profile of said plurality of genes or their encoded proteins to a good and/or poor prognosis template profile of expression levels of said plurality of genes or their encoded proteins, and classifying said individual as having a good prognosis or poor prognosis if said patient expression profile has a high similarity to a good prognosis template or a poor prognosis template, respectively. In a more specific embodiment of the method, said plurality of patient subsets comprises ER⁻, *BRCAl* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high

individuals; ER+, ER/AGE low, LN+ individuals; and ER+, ER/AGE low, LN⁻ individuals. In another embodiment, said good prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a good prognosis or good outcome, while said poor prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a poor prognosis or poor outcome. In another specific embodiment, said good or poor prognosis template is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis, respectively.

[00123] It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining an expression profile of a plurality of the genes for which markers are listed in Tables 1-5 in a sample from said individual; and comparing said expression profile to a template profile, where said template is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER+, ER/AGE high type; ER+, ER/AGE low, LN+ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said template comprises the average levels of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression profile of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a template profile comprising levels of expression of the same genes for each of Tables 1-5, and determining whether the pattern of expression of said genes in said sample is most similar to the pattern of expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

[00124] In another embodiment, the invention provides a method of classifying an individual as having a good prognosis or a poor prognosis, comprising: (a) classifying said individual as

ER⁻, *BRCAl*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining an expression profile of a first plurality of genes in a cell sample taken from the individual relative to a control, said first plurality of genes comprising at least two of the genes corresponding to the markers Table 1 if said individual is classified as ER⁻, *BRCAl*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻, wherein said individual is “ER/AGE high” if the ER level of the individual exceeds a predetermined value, and “ER/AGE low” if the ER level of the individual does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual. In another specific embodiment, said individual is ER⁻, *BRCAl*, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said individual is ER⁻, sporadic, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said individual is ER⁺, ER/AGE high, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

[00125] Where information is available regarding the LN status of a breast cancer patient, the patient may be identified as having a “very good prognosis,” an “intermediate prognosis,” or a poor prognosis, which enables the refinement of treatment. In one embodiment, the invention provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) classifying said patient as having a “poor prognosis,” “intermediate prognosis,” or “very good prognosis” on the basis of the levels of expression of at least five genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5; and (b) assigning said patient a therapeutic regimen, said therapeutic regimen (i) comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a

good prognosis or an intermediate prognosis, or (ii) comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

[00126] In another embodiment, a breast cancer patient is assigned a prognosis by a method comprising (a) determining the breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCAl*; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE low, LN⁻; (c) determining an expression profile comprising at least five genes in a cell sample taken from said breast cancer patient wherein markers for said at least five genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, *BRCAl*; Table 3 if said patient is classified as ER+, ER/AGE high; Table 4 if said patient is classified as ER+, ER/AGE low, LN+; or Table 5 if said patient is classified as ER+, ER/AGE high, LN⁻; (d) determining the similarity of the expression profile of said at least five genes to a template profile comprising levels of expression of said at least five genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific embodiment of the method, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," wherein said breast cancer patient is assigned a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

[00127] The invention also provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) determining the lymph node status for said patient; (b) determining the expression of at least five genes for which markers are listed in Table 5 in a cell sample from said patient, thereby generating an expression profile; (c) classifying said patient as having a "poor prognosis," "intermediate prognosis," or "very good prognosis" on the basis of said expression profile; and (d) assigning said patient a therapeutic regimen, said therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node

negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and classification. In a specific embodiment of this method, said therapeutic regimen assigned to lymph node negative patients classified as having an "intermediate prognosis" additionally comprises adjuvant hormonal therapy. In another specific embodiment of this method, said classifying step (c) is carried out by a method comprising: (a) rank ordering in descending order a plurality of breast cancer tumor samples that compose a pool of breast cancer tumor samples by the degree of similarity between the expression profile of said at least five genes in each of said tumor samples and the expression profile of said at least five genes across all remaining tumor samples that compose said pool, said degree of similarity being expressed as a similarity value; (b) determining an acceptable number of false negatives in said classifying step, wherein a false negative is a breast cancer patient for whom the expression levels of said at least five genes in said cell sample predicts that said breast cancer patient will have no distant metastases within the first five years after initial diagnosis, but who has had a distant metastasis within the first five years after initial diagnosis; (c) determining a similarity value above which in said rank ordered list said acceptable number of tumor samples or fewer are false negatives; (d) selecting said similarity value determined in step (c) as a first threshold similarity value; (e) selecting a second similarity value, greater than said first similarity value, as a second threshold similarity value; and (f) determining the similarity between the expression profile of said at least five genes in a breast cancer tumor sample from the breast cancer patient and the expression profile of said respective at least five genes in said pool, to obtain a patient similarity value, wherein if said patient similarity value equals or exceeds said second threshold similarity value, said patient is classified as having a "very good prognosis"; if said patient similarity value equals or exceeds said first threshold similarity value, but is less than said second threshold similarity value, said patient is classified as having an "intermediate prognosis"; and if said patient similarity value is less than said first threshold similarity value, said patient is classified as having a "poor prognosis." Another specific embodiment of this method comprises determining the estrogen receptor (ER) status of said patient, wherein if said patient is ER positive and lymph node negative, said therapeutic regimen assigned to said patient additionally comprises adjuvant hormonal therapy.

[00128] A patient in any patient subset or clinical class, e.g., any one of the classes described above, can be classified as having a particular prognosis level, e.g., a good prognosis or a poor prognosis, based on the similarity of the patient's cellular constituent profile to an

appropriate template profile for the prognosis level of patients in the clinical class. In one embodiment, a cellular constituent profile corresponding to a certain prognosis level, e.g., a profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of patients having the prognosis level is used as a template for the prognosis level. For example, a good prognosis template profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of good outcome patients or a poor prognosis template profile comprising measurements of the plurality of cellular constituents representative of levels of the cellular constituents in a plurality of poor outcome patients, can be used for determining whether a patient have good or poor prognosis. Here, a good outcome patient is a patient who has non-reoccurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In contrast, a poor outcome patient is a patient who has reoccurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In a preferred embodiment, both periods are 10 years. Tables 1-5 show exemplary template profiles for the respective patient classes. For example, the expression profile of a patient with a combination of ER+, ER/AGE low, LN+ can be compared with the good prognosis template of Table 4 to determine if the patient has good prognosis or poor prognosis.

[00129] The degree of similarity of the patient's cellular constituent profile to a template of a particular prognosis can be used to indicate whether the patient has the particular prognosis. For example, a high degree of similarity indicates that the patient has the particular prognosis, whereas a low degree of similarity indicates that the patient does not have the particular prognosis. In a preferred embodiment, a patient is classified as having a good prognosis profile if the patient's cellular constituent profile has a high similarity to a good prognosis template and/or has a low similarity to a poor prognosis template. In another embodiment, a patient is classified as having a poor prognosis profile if the patient's cellular constituent profile has a low similarity to a good prognosis template and/or has a high similarity to a poor prognosis template. In embodiments for predicting the responsiveness of a breast cancer patient under the age of 55, the patients in the good and poor outcome patient populations used to generate the templates are preferably also under the age of 55 at the time of diagnosis of breast cancer.

[00130] The degree of similarity between a patient's cellular constituent profile and a template profile can be determined using any method known in the art. In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the

template. In one embodiment, a correlation coefficient above a correlation threshold indicates high similarity, whereas a correlation coefficient below the threshold indicates low similarity. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6. In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment, a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

[00131] As an illustration, in one embodiment, a template for a good prognosis is defined as \bar{z}_1 (e.g., a profile consisting of the xdev's listed in the good prognosis column of one of Tables 1-5) and/or a template for poor prognosis is defined as \bar{z}_2 (e.g., a profile consisting of the xdev's listed in the poor prognosis column of one of Tables 1-5). Either one or both of the two classifier parameters (P_1 and P_2) can then be used to measure degrees of similarities between a patient's profile and the respective templates: P_1 measures the similarity between the patient's profile \bar{y} and the good prognosis template \bar{z}_1 , and P_2 measures the similarity between \bar{y} and the poor prognosis template \bar{z}_2 . In embodiments which employ correlation coefficients, the correlation coefficient P_i can be calculated as:

$$P_i = (\bar{z}_i \bullet \bar{y}) / (\|\bar{z}_i\| \cdot \|\bar{y}\|) \quad (4)$$

where $i = 1$ and 2 .

[00132] Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a selected correlation threshold or if P_2 is equal to or less than a selected correlation threshold. In another embodiment, \bar{y} is classified as a poor prognosis profile if P_1 is less than a selected correlation threshold or if P_2 is above a selected correlation threshold. In still another embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a first selected correlation threshold and \bar{y} is classified as a poor prognosis profile if P_2 is greater than a second selected correlation threshold.

[00133] In a preferred embodiment, the cellular constituent profile is an expression profile comprising measurements of a plurality of transcripts (e.g., measured as mRNAs or cDNAs) in a sample derived from a patient, e.g., the plurality of transcripts corresponding to the markers in all or a portion of one of Tables 1-5. In this embodiment, the good prognosis template can be a good prognosis expression template comprising measurements of the

plurality of transcripts representative of expression levels of the transcripts in a plurality of good prognosis patients, and the poor prognosis template can be a poor prognosis expression template comprising measurements of the plurality of transcripts representative of expression levels of the transcripts in a plurality of poor prognosis patients. In a preferred embodiment, measurement of each transcript in the good or poor prognosis expression template is an average of expression levels of the transcript in the plurality of good or poor prognosis patients, respectively.

[00134] In another embodiment, the expression profile is a differential expression profile comprising differential measurements of the plurality of transcripts in a sample derived from the patient versus measurements of the plurality of transcripts in a control sample. The differential measurements can be x_{dev} , $\log(\text{ratio})$, error-weighted $\log(\text{ratio})$, or a mean subtracted $\log(\text{intensity})$ (see, e.g., Stoughton et al., PCT publication WO 00/39339, published on July 6, 2000; U.S. Patent Application No. 10/848,755, filed May 18, 2004, by Mao et al., attorney docket no: 9301-188-999, each of which is incorporated herein by reference in its entirety).

5.4.3 IMPROVING SENSITIVITY TO EXPRESSION LEVEL DIFFERENCES

[00135] In using the markers disclosed herein, and, indeed, using any sets of markers, e.g., to compare profiles or to differentiate an individual having one phenotype from another individual having a second phenotype, one can compare the profile comprising absolute expression levels of the markers in a sample to a template; for example, a template comprising the average levels of expression of the markers in a plurality of individuals. To increase the sensitivity of the comparison, however, the expression level values are preferably transformed in a number of ways. Also, to differentiate an individual having one phenotype from another individual having a second phenotype using any sets of markers, one can compare the absolute expression of each of the markers in a sample to a control; for example, the control can be the average level of expression of each of the markers, respectively, in a pool of individuals.

[00136] For example, the expression level of each of the markers can be normalized by the average expression level of all markers the expression level of which is determined, or by the average expression level of a set of control genes. Thus, in one embodiment, the markers are represented by probes on a microarray, and the expression level of each of the markers is normalized by the mean or median expression level across all of the genes represented on the microarray, including any non-marker genes. In a specific embodiment, the normalization is

carried out by dividing the median or mean level of expression of all of the genes on the microarray. In another embodiment, the expression levels of the markers is normalized by the mean or median level of expression of a set of control markers. In a specific embodiment, the control markers comprise a set of housekeeping genes. In another specific embodiment, the normalization is accomplished by dividing by the median or mean expression level of the control genes.

[00137] The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a pool of samples. Preferably, the comparison is to the mean or median expression level of each the marker genes in the pool of samples. Such a comparison may be accomplished, for example, by dividing by the mean or median expression level of the pool for each of the markers from the expression level each of the markers in the sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the pool as a whole, making comparisons more sensitive and more likely to produce meaningful results that the use of absolute expression levels alone. The expression level data may be transformed in any convenient way; preferably, the expression level data for all is log transformed before means or medians are taken.

[00138] In performing comparisons to a pool, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the pool, where nucleic acid derived from the sample and nucleic acid derived from the pool are hybridized during the course of a single experiment. Such an approach requires that new pool nucleic acid be generated for each comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, and preferably, the expression levels in a pool, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

[00139] The current invention also provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared

value. The first compared value is then compared to a second compared value, wherein said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first compared value is then compared to a third compared value, wherein said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, the marker can be a gene, a protein encoded by the gene, etc. Optionally, the first compared value can be compared to additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional compared values, said first compared value is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said compared value most similar to said first compared value.

[00140] The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a control sample, e.g., a sample comprises a pool of samples, to generate a differential expression profile. Such a comparison may be accomplished, for example, by determining a ratio between expression level of each marker in the sample and the expression level of the corresponding marker in the control sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the control as a whole, making subsequent comparisons to a template more sensitive and more likely to produce meaningful results than the use of absolute expression levels alone. The comparison may be performed in any convenient way, e.g., by taking difference, ratio, or log(ratio).

[00141] In performing comparisons to a control sample, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the control sample, where nucleic acid derived from the sample and nucleic acid derived from the control are hybridized during the course of a single experiment. Such an approach requires that new control sample of nucleic acid be generated for each

comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, the expression levels in a control sample, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

[00142] The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promotor sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 in vitro transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

[00143] The current invention provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms

exhibiting said at least two different phenotypes, respectively, to produce a first compared value so that a first differential profile comprising a plurality of first compared values for said plurality of markers is generated. The first differential profile is then compared to a second differential profile comprising a plurality of second compared values, wherein each said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first differential profile is then compared to a third differential profile comprising a plurality of third compared values, wherein each said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, each marker can be a gene, a protein encoded by the gene, etc. Optionally, the first differential profile can be compared to additional expression profiles each of which comprising additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional differential profiles, said first differential profile is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said differential profile most similar to said first differential profile.

[00144] In a specific embodiment of this method, the compared values are each ratios of the levels of expression of each of said genes. In another specific embodiment, each of the levels of expression of each of the genes in the pooled sample are normalized prior to any of the comparing steps. In a more specific embodiment, the normalization of the levels of expression is carried out by dividing by the median or mean level of the expression of each of the genes or dividing by the mean or median level of expression of one or more housekeeping genes in the pooled sample from said cell or organism. In another specific embodiment, the normalized levels of expression are subjected to a log transform, and the comparing steps comprise subtracting the log transform from the log of the levels of expression of each of the genes in the sample. In another specific embodiment, the two or more different phenotypes are different stages of a disease or disorder. In still another specific embodiment, the two or

more different phenotypes are different prognoses of a disease or disorder. In yet another specific embodiment, the levels of expression of each of the genes, respectively, in the pooled sample or said levels of expression of each of said genes in a sample from the cell or organism characterized as having the first phenotype, second phenotype, or said phenotype different from said first and second phenotypes, respectively, are stored on a computer or on a computer-readable medium.

[00145] In another specific embodiment, the two phenotypes are good prognosis and poor prognosis. In a more specific embodiment, the two phenotypes are good prognosis and poor prognosis for an individual that is identified as having ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status.

[00146] In another specific embodiment, the comparison is made between the expression profile of the genes in the sample and the expression profile of the same genes in a pool representing only one of two or more phenotypes. In the context of prognosis-correlated genes, for example, one can compare the expression levels of prognosis-related genes in a sample to the average levels of the expression of the same genes in a plurality of “good prognosis” samples (as opposed to a plurality of samples that include samples from patients having poor prognoses and good prognoses). Thus, in this method, a sample is classified as having a good prognosis if the expression profile of prognosis-correlated genes exceeds a chosen coefficient of correlation to the average “good prognosis” expression profile (*e.g.*, the profile comprising average levels of expression of prognosis-correlated genes in samples from a plurality of patients having a “good prognosis”). Patients whose expression profiles correlate more poorly with the “good prognosis” expression profile (*e.g.*, whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

[00147] Where individuals are classified on the basis of phenotypic, genotypic, or clinical characteristics into patient subsets, the pool of samples may be a pool of samples for the phenotype that includes samples representing each of the patient subsets. Alternatively, the pool of samples may be a pool of samples for the phenotype representing only the specific patient subset. For example, where an individual is classified as ER⁺, sporadic, the pool of samples to which the individual's sample is compared may be a pool of samples from ER⁺, sporadic individuals having a good prognosis only, or may be a pool of samples of individuals having a good prognosis, without regard to ER status or mutation type.

[00148] The method can be applied to a plurality of patient subsets. For example, in a specific embodiment, the phenotype is good prognosis, and the individual is classified into

one of the following patient subsets: ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status. A set of markers informative for prognosis for the patient subset into which the individual is classified is then used to determine the likely prognosis for the individual. A sample is classified as coming from an individual having a good prognosis if the expression profile of prognosis-correlated genes for the particular subset into which the individual is classified exceeds a chosen coefficient of correlation to the average “good prognosis” expression profile (*e.g.*, the levels of expression of prognosis-correlated genes in a plurality of samples from patients within the subclass having a “good prognosis”). Patients whose expression levels correlate more poorly with the “good prognosis” expression profile (*e.g.*, whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

[00149] Of course, single-channel data may also be used without specific comparison to a mathematical sample pool. For example, a sample may be classified as having a first or a second phenotype, wherein the first and second phenotypes are related, by calculating the similarity between the expression profile of at least 5 markers in the sample, where the markers are correlated with the first or second phenotype, to a first phenotype template and a second phenotype template each comprising the expression levels of the same markers, by (a) labeling nucleic acids derived from a sample with a fluorophore to obtain a pool of fluorophore-labeled nucleic acids; (b) contacting said fluorophore-labeled nucleic acid with a microarray under conditions such that hybridization can occur, detecting at each of a plurality of discrete loci on the microarray a fluorescent emission signal from said fluorophore-labeled nucleic acid that is bound to said microarray under said conditions; and (c) determining the similarity of marker gene expression in the individual sample to the first and second templates, wherein if said expression is more similar to the first template, the sample is classified as having the first phenotype, and if said expression is more similar to the second template, the sample is classified as having the second phenotype.

[0100] In a specific embodiment of the above method, the first phenotype is a good prognosis of breast cancer, the sample is a sample from an individual that has been classified into a patient subset, and the first and second templates are templates for the phenotype for the particular patient subset. In a more specific embodiment, for example, the first phenotype is a good prognosis, the second phenotype is a poor prognosis, the patient is classified into an ER⁻, sporadic patient subset, an ER⁻, *BRCA1* subset, an ER⁺, ER/AGE high subset, an ER⁺, ER/AGE low, LN⁺ subset, or an ER⁺, ER/AGE low, LN⁺ subset, and said first and second

templates are templates derived from the expression of the marker genes in individuals having a good prognosis and a poor prognosis, respectively, wherein said individuals are all of the patient subset into which said patient is classified.

5.5 DETERMINATION OF MARKER GENE EXPRESSION LEVELS

5.5.1 METHODS

[00150] The expression levels of the marker genes in a sample may be determined by any means known in the art. The expression level may be determined by isolating and determining the level (*i.e.*, amount) of nucleic acid transcribed from each marker gene. Alternatively, or additionally, the level of specific proteins encoded by a marker gene may be determined.

[00151] The level of expression of specific marker genes can be accomplished by determining the amount of mRNA, or polynucleotides derived therefrom, present in a sample. Any method for determining RNA levels can be used. For example, RNA is isolated from a sample and separated on an agarose gel. The separated RNA is then transferred to a solid support, such as a filter. Nucleic acid probes representing one or more markers are then hybridized to the filter by northern hybridization, and the amount of marker-derived RNA is determined. Such determination can be visual, or machine-aided, for example, by use of a densitometer. Another method of determining RNA levels is by use of a dot-blot or a slot-blot. In this method, RNA, or nucleic acid derived therefrom, from a sample is labeled. The RNA or nucleic acid derived therefrom is then hybridized to a filter containing oligonucleotides derived from one or more marker genes, wherein the oligonucleotides are placed upon the filter at discrete, easily-identifiable locations. Hybridization, or lack thereof, of the labeled RNA to the filter-bound oligonucleotides is determined visually or by densitometer. Polynucleotides can be labeled using a radiolabel or a fluorescent (*i.e.*, visible) label.

[00152] These examples are not intended to be limiting; other methods of determining RNA abundance are known in the art.

[00153] The level of expression of particular marker genes may also be assessed by determining the level of the specific protein expressed from the marker genes. This can be accomplished, for example, by separation of proteins from a sample on a polyacrylamide gel, followed by identification of specific marker-derived proteins using antibodies in a western blot. Alternatively, proteins can be separated by two-dimensional gel electrophoresis

systems. Two-dimensional gel electrophoresis is well-known in the art and typically involves isoelectric focusing along a first dimension followed by SDS-PAGE electrophoresis along a second dimension. *See, e.g., Hames et al., 1990, GEL ELECTROPHORESIS OF PROTEINS: A PRACTICAL APPROACH, IRL Press, New York; Shevchenko et al., Proc. Nat'l Acad. Sci. USA 93:1440-1445 (1996); Sagliocco et al., Yeast 12:1519-1533 (1996); Lander, Science 274:536-539 (1996).* The resulting electropherograms can be analyzed by numerous techniques, including mass spectrometric techniques, western blotting and immunoblot analysis using polyclonal and monoclonal antibodies.

[00154] Alternatively, marker-derived protein levels can be determined by constructing an antibody microarray in which binding sites comprise immobilized, preferably monoclonal, antibodies specific to a plurality of protein species encoded by the cell genome. Preferably, antibodies are present for a substantial fraction of the marker-derived proteins of interest. Methods for making monoclonal antibodies are well known (*see, e.g., Harlow and Lane, 1988, ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor, New York, which is incorporated in its entirety for all purposes*). In one embodiment, monoclonal antibodies are raised against synthetic peptide fragments designed based on genomic sequence of the cell. With such an antibody array, proteins from the cell are contacted to the array, and their binding is assayed with assays known in the art. Generally, the expression, and the level of expression, of proteins of diagnostic or prognostic interest can be detected through immunohistochemical staining of tissue slices or sections.

[00155] Finally, expression of marker genes in a number of tissue specimens may be characterized using a "tissue array" (Kononen *et al.*, *Nat. Med* 4(7):844-7 (1998)). In a tissue array, multiple tissue samples are assessed on the same microarray. The arrays allow *in situ* detection of RNA and protein levels; consecutive sections allow the analysis of multiple samples simultaneously.

5.5.2 MICROARRAYS

[00156] In preferred embodiments, polynucleotide microarrays are used to measure expression so that the expression status of each of the markers above is assessed simultaneously. Generally, microarrays according to the invention comprise a plurality of markers informative for prognosis, or outcome determination, for a particular disease or condition, and, in particular, for individuals having specific combinations of genotypic or phenotypic characteristics of the disease or condition (*i.e.*, that are prognosis-informative for a particular patient subset).

[00157] The microarrays of the invention preferably comprise at least 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more of markers, or all of the markers, or any combination of markers, identified as prognosis-informative within a patient subset. The actual number of informative markers the microarray comprises will vary depending upon the particular condition of interest, the number of markers identified, and, optionally, the number of informative markers found to result in the least Type I error, Type II error, or Type I and Type II error in determination of prognosis. As used herein, "Type I error" means a false positive and "Type II error" means a false negative; in the example of prognosis of breast cancer, Type I error is the mis-characterization of an individual with a good prognosis as having a poor prognosis, and Type II error is the mis-characterization of an individual with a poor prognosis as having a good prognosis.

[00158] In specific embodiments, the invention provides polynucleotide arrays in which the prognosis markers identified for a particular patient subset comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for a particular patient subset. Microarrays of the invention, of course, may comprise probes complementary and hybridizable to prognosis-informative markers for a plurality of the patient subsets, or for each patient subset, identified for a particular condition. In another embodiment, therefore, the microarray of the invention comprises a plurality of probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for each patient subset identified for the condition of interest, and wherein said probes, in total, are at least 50% of the probes on said microarray.

[00159] In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers identified by the methods disclosed elsewhere herein. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers identified for the patient subset(s) for the condition of interest, and can be used to monitor marker expression in the same way as a microarray containing only prognosis-informative markers otherwise identified.

[00160] In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five markers identified by the methods described herein. Preferably, a commercially-available cDNA microarray comprises all of the markers identified by the methods described herein as being informative for a patient subset for a particular condition. However, such a microarray may comprise at least 5, 10, 15 or 25 of such markers, up to the maximum number of markers identified.

[00161] In an embodiment specific to breast cancer, the invention provides for oligonucleotide or cDNA arrays comprising probes hybridizable to the genes corresponding to each of the marker sets described above (*i.e.*, markers informative for ER⁻, sporadic individuals, markers informative for ER⁻, *BRCAl* individuals, markers informative for ER⁺, ER/AGE high individuals, markers informative for ER⁺, ER/AGE low, LN⁺ individuals, and markers informative for ER⁺, ER/AGE low, LN⁻ individuals, as shown in Tables 1-5). Any of the microarrays described herein may be provided in a sealed container in a kit.

[00162] The invention provides microarrays containing probes useful for the prognosis of any breast cancer patient, or for breast cancer patients classified into one of a plurality of patient subsets. In particular, the invention provides polynucleotide arrays comprising probes to a subset or subsets of at least 5, 10, 15, 20, 25 or more of the genetic markers, or up to the full set of markers, in any of Tables 1-5, which distinguish between patients with good and poor prognosis. In certain embodiments, therefore, the invention provides microarrays comprising probes for a plurality of the genes for which markers are listed in Tables 1, 2, 3, 4 or 5. In a specific embodiment, the microarray of the invention comprises 1, 2, 3, 4, 5 or 10 of the markers in Table 1, at least five of the markers in Table 2; 1, 2, 3, 4, 5 or 10 of the markers in Table 3; 1, 2, 3, 4, 5 or 10 of the markers in Table 4; or 1, 2, 3, 4, 5 or 10 of the markers in Table 1. In other embodiments, the microarray comprises probes for 1, 2, 3, 4, 5, or 10 of the markers shown in any two, three or four of Tables 1-5, or all of Tables 1-5. In other embodiments, the microarray of the invention contains each of the markers in Table 1, Table 2, Table 3, Table 4, or Table 5. In another embodiment, the microarray contains all of the markers shown in Tables 1-5. In specific embodiments, the array comprises probes derived only from the markers listed in Table 1, Table 2, Table 3, Table 4, or Table 5; probes derived from any two of Tables 1-5; any three of Tables 1-5; any four of Tables 1-5; or all of Tables 1-5.

[00163] In other embodiments, the array comprises a plurality of probes derived from markers listed in any of Tables 1-5 in combination with a plurality of other probes, derived

from markers not listed in any of Tables 1-5, that are identified as informative for the prognosis of breast cancer.

[00164] In specific embodiments, the invention provides polynucleotide arrays in which the breast cancer prognosis markers described herein in Tables 1, 2, 3, 4 and/or 5 comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

[00165] In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers listed in Tables 1-5. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers listed in Tables 1-5, or a subset thereof as described above, and can be used to monitor marker expression in the same way as a microarray containing only markers listed in Tables 1-5.

[00166] In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five of the markers listed in Tables 1-5. Preferably, a commercially-available cDNA microarray comprises all of the markers listed in Tables 1-5. However, such a microarray may comprise at least 5, 10, 15 or 25 of the markers in any of Tables 1-5, up to the maximum number of markers in a Table, and may comprise all of the markers in any one of Tables 1-5, and a subset of another of Tables 1-5, or subsets of each as described above. In a specific embodiment of the microarrays used in the methods disclosed herein, the markers that are all or a portion of Tables 1-5 make up at least 50%, 60%, 70%, 80%, 90%, 95% or 98% of the probes on the microarray.

[00167] General methods pertaining to the construction of microarrays comprising the marker sets and/or subsets above are described in the following sections.

[00168]

[00169]

5.5.2.1 CONSTRUCTION OF MICROARRAYS

[00170] Microarrays are prepared by selecting probes which comprise a polynucleotide sequence, and then immobilizing such probes to a solid support or surface. For example, the probes may comprise DNA sequences, RNA sequences, or copolymer sequences of DNA and RNA. The polynucleotide sequences of the probes may also comprise DNA and/or RNA analogues, or combinations thereof. For example, the polynucleotide sequences of the probes may be full or partial fragments of genomic DNA. The polynucleotide sequences of the probes may also be synthesized nucleotide sequences, such as synthetic oligonucleotide sequences. The probe sequences can be synthesized either enzymatically *in vivo*, enzymatically *in vitro* (e.g., by PCR), or non-enzymatically *in vitro*.

[00171] The probe or probes used in the methods of the invention are preferably immobilized to a solid support which may be either porous or non-porous. For example, the probes of the invention may be polynucleotide sequences which are attached to a nitrocellulose or nylon membrane or filter covalently at either the 3' or the 5' end of the polynucleotide. Such hybridization probes are well known in the art (see, e.g., Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). Alternatively, the solid support or surface may be a glass or plastic surface. In a particularly preferred embodiment, hybridization levels are measured to microarrays of probes consisting of a solid phase on the surface of which are immobilized a population of polynucleotides, such as a population of DNA or DNA mimics, or, alternatively, a population of RNA or RNA mimics. The solid phase may be a nonporous or, optionally, a porous material such as a gel.

[00172] In preferred embodiments, a microarray comprises a support or surface with an ordered array of binding (e.g., hybridization) sites or "probes" each representing one of the markers described herein. Preferably the microarrays are addressable arrays, and more preferably positionally addressable arrays. More specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position in the array

(*i.e.*, on the support or surface). In preferred embodiments, each probe is covalently attached to the solid support at a single site.

[00173] Microarrays can be made in a number of ways, of which several are described below. However produced, microarrays share certain characteristics. The arrays are reproducible, allowing multiple copies of a given array to be produced and easily compared with each other. Preferably, microarrays are made from materials that are stable under binding (*e.g.*, nucleic acid hybridization) conditions. The microarrays are preferably small, *e.g.*, between 1 cm² and 25 cm², between 12 cm² and 13 cm², or 3 cm². However, larger arrays are also contemplated and may be preferable, *e.g.*, for use in screening arrays. Preferably, a given binding site or unique set of binding sites in the microarray will specifically bind (*e.g.*, hybridize) to the product of a single gene in a cell (*e.g.*, to a specific mRNA, or to a specific cDNA derived therefrom). However, in general, other related or similar sequences will cross hybridize to a given binding site.

[00174] The microarrays of the present invention include one or more test probes, each of which has a polynucleotide sequence that is complementary to a subsequence of RNA or DNA to be detected. Preferably, the position of each probe on the solid surface is known. Indeed, the microarrays are preferably positionally addressable arrays. Specifically, each probe of the array is preferably located at a known, predetermined position on the solid support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position on the array (*i.e.*, on the support or surface).

[00175] According to the invention, the microarray is an array (*i.e.*, a matrix) in which each position represents one of the markers described herein. For example, each position can contain a DNA or DNA analogue based on genomic DNA to which a particular RNA or cDNA transcribed from that genetic marker can specifically hybridize. The DNA or DNA analogue can be, *e.g.*, a synthetic oligomer or a gene fragment. In one embodiment, probes representing each of the markers is present on the array. In a preferred embodiment, the array comprises probes for each of the markers listed in Tables 1-5.

5.5.2.2 PREPARING PROBES FOR MICROARRAYS

[00176] As noted above, the "probe" to which a particular polynucleotide molecule specifically hybridizes according to the invention contains a complementary genomic polynucleotide sequence. The probes of the microarray preferably consist of nucleotide sequences of no more than 1,000 nucleotides. In some embodiments, the probes of the array consist of nucleotide sequences of 10 to 1,000 nucleotides. In a preferred embodiment, the

nucleotide sequences of the probes are in the range of 10-200 nucleotides in length and are genomic sequences of a species of organism, such that a plurality of different probes is present, with sequences complementary and thus capable of hybridizing to the genome of such a species of organism, sequentially tiled across all or a portion of such genome. In other specific embodiments, the probes are in the range of 10-30 nucleotides in length, in the range of 10-40 nucleotides in length, in the range of 20-50 nucleotides in length, in the range of 40-80 nucleotides in length, in the range of 50-150 nucleotides in length, in the range of 80-120 nucleotides in length, and most preferably are 60 nucleotides in length.

[00177] The probes may comprise DNA or DNA “mimics” (e.g., derivatives and analogues) corresponding to a portion of an organism’s genome. In another embodiment, the probes of the microarray are complementary RNA or RNA mimics. DNA mimics are polymers composed of subunits capable of specific, Watson-Crick-like hybridization with DNA, or of specific hybridization with RNA. The nucleic acids can be modified at the base moiety, at the sugar moiety, or at the phosphate backbone. Exemplary DNA mimics include, e.g., phosphorothioates.

[00178] DNA can be obtained, e.g., by polymerase chain reaction (PCR) amplification of genomic DNA or cloned sequences. PCR primers are preferably chosen based on a known sequence of the genome that will result in amplification of specific fragments of genomic DNA. Computer programs that are well known in the art are useful in the design of primers with the required specificity and optimal amplification properties, such as *Oligo* version 5.0 (National Biosciences). Typically each probe on the microarray will be between 10 bases and 50,000 bases, usually between 300 bases and 1,000 bases in length. PCR methods are well known in the art, and are described, for example, in Innis *et al.*, eds., *PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS*, Academic Press Inc., San Diego, CA (1990). It will be apparent to one skilled in the art that controlled robotic systems are useful for isolating and amplifying nucleic acids.

[00179] An alternative, preferred means for generating the polynucleotide probes of the microarray is by synthesis of synthetic polynucleotides or oligonucleotides, e.g., using N-phosphonate or phosphoramidite chemistries (Froehler *et al.*, *Nucleic Acid Res.* 14:5399-5407 (1986); McBride *et al.*, *Tetrahedron Lett.* 24:246-248 (1983)). Synthetic sequences are typically between about 10 and about 500 bases in length, more typically between about 20 and about 100 bases, and most preferably between about 40 and about 70 bases in length. In some embodiments, synthetic nucleic acids include non-natural bases, such as, but by no means limited to, inosine. As noted above, nucleic acid analogues may be used as binding

sites for hybridization. An example of a suitable nucleic acid analogue is peptide nucleic acid (see, e.g., Egholm *et al.*, *Nature* 363:566-568 (1993); U.S. Patent No. 5,539,083).

[00180] Probes are preferably selected using an algorithm that takes into account binding energies, base composition, sequence complexity, cross-hybridization binding energies, and secondary structure. See Friend *et al.*, International Patent Publication WO 01/05935, published January 25, 2001; Hughes *et al.*, *Nat. Biotech.* 19:342-7 (2001).

[00181] A skilled artisan will also appreciate that positive control probes, e.g., probes known to be complementary and hybridizable to sequences in the target polynucleotide molecules, and negative control probes, e.g., probes known to not be complementary and hybridizable to sequences in the target polynucleotide molecules, should be included on the array. In one embodiment, positive controls are synthesized along the perimeter of the array. In another embodiment, positive controls are synthesized in diagonal stripes across the array. In still another embodiment, the reverse complement for each probe is synthesized next to the position of the probe to serve as a negative control. In yet another embodiment, sequences from other species of organism are used as negative controls or as "spike-in" controls.

5.5.2.3 ATTACHING PROBES TO THE SOLID SURFACE

[00182] The probes are attached to a solid support or surface, which may be made, e.g., from glass, plastic (e.g., polypropylene, nylon), polyacrylamide, nitrocellulose, gel, or other porous or nonporous material. A preferred method for attaching the nucleic acids to a surface is by printing on glass plates, as is described generally by Schena *et al.*, *Science* 270:467-470 (1995). This method is especially useful for preparing microarrays of cDNA (See also, DeRisi *et al.*, *Nature Genetics* 14:457-460 (1996); Shalon *et al.*, *Genome Res.* 6:639-645 (1996); and Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10539-11286 (1995)).

[00183] A second preferred method for making microarrays is by making high-density oligonucleotide arrays. Techniques are known for producing arrays containing thousands of oligonucleotides complementary to defined sequences, at defined locations on a surface using photolithographic techniques for synthesis *in situ* (see, Fodor *et al.*, 1991, *Science* 251:767-773; Pease *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:5022-5026; Lockhart *et al.*, 1996, *Nature Biotechnology* 14:1675; U.S. Patent Nos. 5,578,832; 5,556,752; and 5,510,270) or other methods for rapid synthesis and deposition of defined oligonucleotides (Blanchard *et al.*, *Biosensors & Bioelectronics* 11:687-690). When these methods are used, oligonucleotides (e.g., 60-mers) of known sequence are synthesized directly on a surface such

as a derivatized glass slide. Usually, the array produced is redundant, with several oligonucleotide molecules per RNA.

[00184] Other methods for making microarrays, *e.g.*, by masking (Maskos and Southern, 1992, *Nuc. Acids. Res.* 20:1679-1684), may also be used. In principle, and as noted *supra*, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) could be used. However, as will be recognized by those skilled in the art, very small arrays will frequently be preferred because hybridization volumes will be smaller.

[00185] In one embodiment, the arrays of the present invention are prepared by synthesizing polynucleotide probes on a support. In such an embodiment, polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

[00186] In a particularly preferred embodiment, microarrays of the invention are manufactured by means of an ink jet printing device for oligonucleotide synthesis, *e.g.*, using the methods and systems described by Blanchard in U.S. Pat. No. 6,028,189; Blanchard *et al.*, 1996, *Biosensors and Bioelectronics* 11:687-690; Blanchard, 1998, in *Synthetic DNA Arrays in Genetic Engineering*, Vol. 20, J.K. Setlow, Ed., Plenum Press, New York at pages 111-123. Specifically, the oligonucleotide probes in such microarrays are preferably synthesized in arrays, *e.g.*, on a glass slide, by serially depositing individual nucleotide bases in "microdroplets" of a high surface tension solvent such as propylene carbonate. The microdroplets have small volumes (*e.g.*, 100 pL or less, more preferably 50 pL or less) and are separated from each other on the microarray (*e.g.*, by hydrophobic domains) to form circular surface tension wells which define the locations of the array elements (*i.e.*, the different probes). Microarrays manufactured by this ink-jet method are typically of high density, preferably having a density of at least about 2,500 different probes per 1 cm². The polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

5.5.2.4 TARGET POLYNUCLEOTIDE MOLECULES

[00187] The polynucleotide molecules which may be analyzed by the present invention (the "target polynucleotide molecules") may be from any clinically relevant source, but are expressed RNA or a nucleic acid derived therefrom (*e.g.*, cDNA or amplified RNA derived from cDNA that incorporates an RNA polymerase promoter), including naturally occurring nucleic acid molecules, as well as synthetic nucleic acid molecules. In one embodiment, the

target polynucleotide molecules comprise RNA, including, but by no means limited to, total cellular RNA, poly(A)⁺ messenger RNA (mRNA) or fraction thereof, cytoplasmic mRNA, or RNA transcribed from cDNA (*i.e.*, cRNA; see, *e.g.*, Linsley & Schelter, U.S. Patent Application No. 09/411,074, filed October 4, 1999, or U.S. Patent Nos. 5,545,522, 5,891,636, or 5,716,785). Methods for preparing total and poly(A)⁺ RNA are well known in the art, and are described generally, *e.g.*, in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). In one embodiment, RNA is extracted from cells of the various types of interest in this invention using guanidinium thiocyanate lysis followed by CsCl centrifugation (Chirgwin *et al.*, 1979, *Biochemistry* 18:5294-5299). In another embodiment, total RNA is extracted using a silica gel-based column, commercially available examples of which include RNeasy (Qiagen, Valencia, California) and StrataPrep (Stratagene, La Jolla, California). In an alternative embodiment, which is preferred for *S. cerevisiae*, RNA is extracted from cells using phenol and chloroform, as described in Ausubel *et al.*, eds., 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Vol. III, Green Publishing Associates, Inc., John Wiley & Sons, Inc., New York, at pp. 13.12.1-13.12.5). Poly(A)⁺ RNA can be selected, *e.g.*, by selection with oligo-dT cellulose or, alternatively, by oligo-dT primed reverse transcription of total cellular RNA. In one embodiment, RNA can be fragmented by methods known in the art, *e.g.*, by incubation with ZnCl₂, to generate fragments of RNA. In another embodiment, the polynucleotide molecules analyzed by the invention comprise cDNA, or PCR products of amplified RNA or cDNA.

[00188] In one embodiment, total RNA, mRNA, or nucleic acids derived therefrom, is isolated from a sample taken from a person afflicted with breast cancer. Target polynucleotide molecules that are poorly expressed in particular cells may be enriched using normalization techniques (Bonaldi *et al.*, 1996, *Genome Res.* 6:791-806).

[00189] As described above, the target polynucleotides are detectably labeled at one or more nucleotides. Any method known in the art may be used to detectably label the target polynucleotides. Preferably, this labeling incorporates the label uniformly along the length of the RNA, and more preferably, the labeling is carried out at a high degree of efficiency. One embodiment for this labeling uses oligo-dT primed reverse transcription to incorporate the label; however, conventional methods of this method are biased toward generating 3' end fragments. Thus, in a preferred embodiment, random primers (*e.g.*, 9-mers) are used in reverse transcription to uniformly incorporate labeled nucleotides over the full length of the

target polynucleotides. Alternatively, random primers may be used in conjunction with PCR methods or T7 promoter-based *in vitro* transcription methods in order to amplify the target polynucleotides.

[00190] In a preferred embodiment, the detectable label is a luminescent label. For example, fluorescent labels, bioluminescent labels, chemiluminescent labels, and colorimetric labels may be used in the present invention. In a highly preferred embodiment, the label is a fluorescent label, such as a fluorescein, a phosphor, a rhodamine, or a polymethine dye derivative. Examples of commercially available fluorescent labels include, for example, fluorescent phosphoramidites such as FluorePrime (Amersham Pharmacia, Piscataway, N.J.), FluoreDite (Millipore, Bedford, Mass.), FAM (ABI, Foster City, Calif.), and Cy3 or Cy5 (Amersham Pharmacia, Piscataway, N.J.). In another embodiment, the detectable label is a radiolabeled nucleotide.

[00191] In a further preferred embodiment, target polynucleotide molecules from a patient sample are labeled differentially from target polynucleotide molecules of a standard. The standard can comprise target polynucleotide molecules from normal individuals (*i.e.*, those not afflicted with breast cancer). In a highly preferred embodiment, the standard comprises target polynucleotide molecules pooled from samples from normal individuals or tumor samples from individuals having sporadic-type breast tumors. In another embodiment, the target polynucleotide molecules are derived from the same individual, but are taken at different time points, and thus indicate the efficacy of a treatment by a change in expression of the markers, or lack thereof, during and after the course of treatment (*i.e.*, chemotherapy, radiation therapy or cryotherapy), wherein a change in the expression of the markers from a poor prognosis pattern to a good prognosis pattern indicates that the treatment is efficacious. In this embodiment, different timepoints are differentially labeled.

5.5.2.5 HYBRIDIZATION TO MICROARRAYS

[00192] Nucleic acid hybridization and wash conditions are chosen so that the target polynucleotide molecules specifically bind or specifically hybridize to the complementary polynucleotide sequences of the array, preferably to a specific array site, wherein its complementary DNA is located.

[00193] Arrays containing double-stranded probe DNA situated thereon are preferably subjected to denaturing conditions to render the DNA single-stranded prior to contacting with the target polynucleotide molecules. Arrays containing single-stranded probe DNA (*e.g.*, synthetic oligodeoxyribonucleic acids) may need to be denatured prior to contacting with the

target polynucleotide molecules, *e.g.*, to remove hairpins or dimers which form due to self complementary sequences.

[00194] Optimal hybridization conditions will depend on the length (*e.g.*, oligomer versus polynucleotide greater than 200 bases) and type (*e.g.*, RNA, or DNA) of probe and target nucleic acids. One of skill in the art will appreciate that as the oligonucleotides become shorter, it may become necessary to adjust their length to achieve a relatively uniform melting temperature for satisfactory hybridization results. General parameters for specific (*i.e.*, stringent) hybridization conditions for nucleic acids are described in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989), and in Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994). Typical hybridization conditions for the cDNA microarrays of Schena *et al.* are hybridization in 5 X SSC plus 0.2% SDS at 65°C for four hours, followed by washes at 25°C in low stringency wash buffer (1 X SSC plus 0.2% SDS), followed by 10 minutes at 25°C in higher stringency wash buffer (0.1 X SSC plus 0.2% SDS) (Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10614 (1993)). Useful hybridization conditions are also provided in, *e.g.*, Tijessen, 1993, HYBRIDIZATION WITH NUCLEIC ACID PROBES, Elsevier Science Publishers B.V.; and Kricka, 1992, NONISOTOPIC DNA PROBE TECHNIQUES, Academic Press, San Diego, CA.

[00195] Particularly preferred hybridization conditions include hybridization at a temperature at or near the mean melting temperature of the probes (*e.g.*, within 51°C, more preferably within 21°C) in 1 M NaCl, 50 mM MES buffer (pH 6.5), 0.5% sodium sarcosine and 30% formamide.

5.5.2.6 SIGNAL DETECTION AND DATA ANALYSIS

[00196] When fluorescently labeled probes are used, the fluorescence emissions at each site of a microarray may be, preferably, detected by scanning confocal laser microscopy. In one embodiment, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Alternatively, a laser may be used that allows simultaneous specimen illumination at wavelengths specific to the two fluorophores and emissions from the two fluorophores can be analyzed simultaneously (*see* Shalon *et al.*, 1996, "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization," *Genome Research* 6:639-645, which is incorporated by reference in its entirety for all purposes). In a preferred embodiment, the arrays are scanned with a laser

fluorescent scanner with a computer controlled X-Y stage and a microscope objective. Sequential excitation of the two fluorophores is achieved with a multi-line, mixed gas laser and the emitted light is split by wavelength and detected with two photomultiplier tubes. Fluorescence laser scanning devices are described in Schena *et al.*, *Genome Res.* 6:639-645 (1996), and in other references cited herein. Alternatively, the fiber-optic bundle described by Ferguson *et al.*, *Nature Biotech.* 14:1681-1684 (1996), may be used to monitor mRNA abundance levels at a large number of sites simultaneously.

[00197] Signals are recorded and, in a preferred embodiment, analyzed by computer, *e.g.*, using a 12 or 16 bit analog to digital board. In one embodiment the scanned image is despeckled using a graphics program (*e.g.*, Hijaak Graphics Suite) and then analyzed using an image gridding program that creates a spreadsheet of the average hybridization at each wavelength at each site. If necessary, an experimentally determined correction for "cross talk" (or overlap) between the channels for the two fluors may be made. For any particular hybridization site on the transcript array, a ratio of the emission of the two fluorophores can be calculated. The ratio is independent of the absolute expression level of the cognate gene, but is useful for genes whose expression is significantly modulated in association with the different breast cancer-related condition.

5.6 THERAPEUTIC REGIMENS SPECIFIC TO PATIENT SUBSETS

[00198] The benefit of identifying subsets of individuals that have a common condition, followed by identification of sets of genes informative for those particular subsets of individuals, is that such subdivision and identification tends to more accurately identify the subset of genes responsible for, or most closely associated with, a particular form of the condition. For example, breast cancer is a complex condition brought about by several different molecular mechanisms. ER+ individuals, particularly ER+, ER/AGE high individuals, show an increased level of expression of cell cycle-control genes, and the expression of these genes is highly informative for prognosis in this patient subset (*see* Examples). In ER⁻ individuals, however, the expression of these genes is not informative for prognosis.

[00199] The set of informative markers, therefore, can be used to assign a particular course of therapy to an individual, *e.g.*, an individual having breast cancer, depending upon the condition subset into which the individual is classified. In one embodiment, therefore, the invention provides a method of assigning a course of therapy to an individual having a condition, said method comprising classifying the individual into one of a plurality of subsets

of a condition, wherein a plurality of informative genes has been identified for at least one of said subsets; and assigning a course of therapy known or suspected to be effective for treating the subset of the condition associated with those genes. In a specific embodiment, said condition is breast cancer, said patient subset is ER+, ER/AGE high status, and said course of therapy comprises the administration of one or more compounds known or suspected to be effective at arresting the cell cycle. In a more specific embodiment, said one or more compounds comprises taxol or a vinca alkaloid.

[00200] Of course, any course of therapy selected or assigned on the basis of the above phenotypes and gene expression may be supplemented by other treatments or courses of therapy relevant to or known or suspected to be effective in the treatment of the condition. For example, the treatment of breast cancer may additionally comprise surgery, either tissue-preserving or radical, radiation treatment, chemotherapy other than that suggested by gene expression analysis, or any other therapy or treatment known or suspected to be effective.

5.7 CLINICAL TRIALS AND EPIDEMIOLOGICAL STUDIES

[00201] The method of the present invention may also be used to assign individuals to categories within a clinical trial, epidemiological study or the like. For example, individuals may be distinguished according to a characteristic of a condition, such as the presence or absence of specific proteins (*e.g.*, estrogen receptor) or tissue structures (*e.g.*, lymph nodes), and with prognosis, and the results of the trial correlated with prognosis. In a specific example, the condition is breast cancer, the characteristic is the presence of the estrogen receptor, and the outcome is prognosis is the expected reoccurrence or non-reoccurrence of metastases within a given period, for example, five years, after initial diagnosis. In another specific example, the condition is obesity, the characteristics are 24-hour energy expenditure, and the prognosis is the expected occurrence of heart disease or diabetes. In another specific example, the condition is a neurodegenerative disease, the characteristic is exposure to a particular range of concentration of an environmental toxin, and the prognosis is expected occurrence or degree of loss of motor function. In each case, the characteristics and expected outcome are used to assign the individual to a category within a clinical trial or epidemiological study.

[00202] Thus, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; determining the level of expression, in a sample derived from

said individual, of a plurality of genes informative for said condition category; determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and assigning the individual to a category in a clinical trial on the basis of prognosis.

[00203] In a specific embodiment, the invention provides a method of assigning an individual to a category in a breast cancer clinical trial, said method comprising: (a) classifying said individual as ER⁻, *BRCAl*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCAl*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to at least one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another more specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

5.8 KITS

[00204] The present invention further provides for kits comprising the marker sets described above. The components of the kits of the present invention are preferably contained in sealed containers. In a preferred embodiment, the kit comprises a microarray ready for hybridization to target polynucleotide molecules. In specific embodiments, the kit may comprise any of the microarrays described in detail in Section 5.5.2. Where proteins are the target molecules, the kit preferably comprises a plurality of antibodies for binding to specific

condition-related proteins, and means for identifying such binding (*e.g.*, means for performing a sandwich assay, ELISA, RIA, *etc.*). Such antibodies may be provided, for example, individually or as part of an antibody array. The kit may additionally comprise software for the data analyses described above, as described in detail in Section 5.9. The kit preferably contains one or more control samples. Such a control sample may be an artificial population of marker-related or marker-derived polynucleotides suitable for hybridization to a microarray, wherein the markers are related to or relevant to the condition of interest (for example, breast cancer). The control may also, or alternatively, be a set of expression values stored on a computer disk or other storage medium.

[00205] The kits of the invention may be primarily diagnostic in nature; that is, they may assist a physician or researcher in determining a characteristic, for example, the prognosis, of a condition of interest, the likely response to a therapeutic regimen, the likely outcome of exposure to an environmental condition, such as toxin exposure, *etc.* The kits of the invention may also be used to classify individuals, for example, to place individuals into different groups in a clinical trial. The use of each kit is determined by the markers, microarrays, controls, *etc.* included.

[00206] COMPUTER-FACILITATED ANALYSIS The analytic methods described in the previous sections can be implemented by use of the following computer systems and according to the following programs and methods. A computer system comprises internal components linked to external components. The internal components of a typical computer system include a processor element interconnected with a main memory. For example, the computer system can be based on an Intel 8086-, 80386-, 80486-, Pentium™, or Pentium™-based processor with preferably 32 MB or more of main memory. The computer system may also be a Macintosh or a Macintosh-based system, but may also be a minicomputer or mainframe.

[00207] The external components preferably include mass storage. This mass storage can be one or more hard disks (which are typically packaged together with the processor and memory). Such hard disks are preferably of 1 GB or greater storage capacity. Other external components include a user interface device, which can be a monitor, together with an inputting device, which can be a "mouse", or other graphic input devices, and/or a keyboard. A printing device can also be attached to the computer.

[00208] Typically, a computer system is also linked to network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area

communication networks, such as the Internet. This network link allows the computer system to share data and processing tasks with other computer systems.

[00209] Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant invention. These software components collectively cause the computer system to function according to the methods of this invention. These software components are typically stored on the mass storage device. A software component comprises the operating system, which is responsible for managing computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows[®] family, such as Windows 3.1, Windows 95, Windows 98, Windows 2000, or Windows NT, or may be of the Macintosh OS family, or may be UNIX, a UNIX derivative such as LINUX, or an operating system specific to a minicomputer or mainframe. The software component represents common languages and functions conveniently present on this system to assist programs implementing the methods specific to this invention. Many high or low level computer languages can be used to program the analytic methods of this invention. Instructions can be interpreted during run-time or compiled. Preferred languages include C/C++, FORTRAN and JAVA. Most preferably, the methods of this invention are programmed in mathematical software packages that allow symbolic entry of equations and high-level specification of processing, including some or all of the algorithms to be used, thereby freeing a user of the need to procedurally program individual equations or algorithms. Such packages include Matlab from Mathworks (Natick, MA), Mathematica[®] from Wolfram Research (Champaign, IL), or S-Plus[®] from Math Soft (Cambridge, MA). Specifically, the software component includes the analytic methods of the invention as programmed in a procedural language or symbolic package.

[00210] The software to be included with the kit comprises the data analysis methods of the invention as disclosed herein. In particular, the software may include mathematical routines for marker discovery, including the calculation of similarity values between clinical categories (e.g., prognosis) and marker expression. The software may also include mathematical routines for calculating the similarity between sample marker expression and control marker expression, using array-generated fluorescence data, to determine the clinical classification of a sample.

[00211] Additionally, the software may also include mathematical routines for determining the prognostic outcome, and recommended therapeutic regimen, for an individual with a condition of interest. In the specific example of breast cancer, the mathematical routines

would determine the prognostic outcome and recommended therapeutic regimen for an individual having breast cancer. Such breast cancer-specific software would include instructions for the computer system's processor to receive data structures that include the level of expression of five or more of the marker genes listed in any of Tables 1-5 in a breast cancer tumor sample obtained from the breast cancer patient; the mean level of expression of the same genes in a control or template; and the breast cancer patient's clinical information, including age, lymph node status and ER status. The software may additionally include mathematical routines for transforming the hybridization data and for calculating the similarity between the expression levels for the marker genes in the patient's breast cancer tumor sample and a control or template. In a specific embodiment, the software includes mathematical routines for calculating a similarity metric, such as a coefficient of correlation, representing the similarity between the expression levels for the marker genes in the patient's breast cancer tumor sample and the control or template, and expressing the similarity as that similarity metric.

[00212] The software preferably would include decisional routines that integrate the patient's clinical and marker gene expression data, and recommend a course of therapy. In one embodiment, for example, the software causes the processor unit to receive expression data for prognosis-related genes in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a template or control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In a specific example, the software additionally causes the processor unit to receive data structures comprising clinical information about the breast cancer patient. In a more specific example, such clinical information includes the patient's age, estrogen receptor status, and lymph node status.

[00213] The software preferably causes the processor unit to receive data structures comprising relevant phenotypic and/or genotypic characteristics of the particular condition of interest, and/or of an individual having that condition, and classifies the individual into a condition subset according to those characteristics. The software then causes the processor to receive values for subset-specific markers, to calculate a metric of similarity of the values associated with those markers (*e.g.*, level, abundance, activity, *etc.*) from the individual to a control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In the

specific example of breast cancer and a breast cancer patient, the software, in one embodiment, causes the processor unit to receive data structures comprising the patient's age, estrogen receptor status, and lymph node status, and on the basis of this data, to classify the patient into one of the following patient subsets: ER^- , sporadic; ER^- , *BRCA1*; ER^+ , AR/AGE high; ER^+ , ER/AGE low, LN^+ ; or ER^+ , ER/AGE low, LN^- . The software then causes the processor to receive expression values for subset-specific prognosis-informative gene expression in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a patient subset-specific template or control, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen.

[00214] Where the control is an expression template comprising expression values for marker genes within a group of patients, *e.g.*, breast cancer patients, the control can comprise either hybridization data obtained at the same time (*i.e.*, in the same hybridization experiment) as the patient's individual hybridization data, or can be a set of hybridization or marker expression values stored on a computer, or on computer-readable media. If the latter is used, new patient hybridization data for the selected marker genes, obtained from initial or follow-up tumor samples, or suspected tumor samples, can be compared to the stored values for the same genes without the need for additional control hybridizations. However, the software may additionally comprise routines for updating the control data set, *e.g.*, to add information from additional breast cancer patients or to remove existing members of the control data set, and, consequently, for recalculating the average expression level values that comprise the template. In another specific embodiment, said control comprises a set of single-channel mean hybridization intensity values for each of said at least five of said genes, stored on a computer-readable medium.

[00215] Clinical data relating to a breast cancer patient, or a patient having another type of condition, and used by the computer program products of the invention, can be contained in a database of clinical data in which information on each patient is maintained in a separate record, which record may contain any information relevant to the patient, the patient's medical history, treatment, prognosis, or participation in a clinical trial or study, including expression profile data generated as part of an initial diagnosis or for tracking the progress of the condition, for example, breast cancer, during treatment.

[00216] Thus, one embodiment of the invention provides a computer program product for classifying a breast cancer patient according to prognosis, the computer program product for

use in conjunction with a computer having a memory and a processor, the computer program product comprising a computer readable storage medium having a computer program mechanism encoded thereon, wherein said computer program product can be loaded into the one or more memory units of a computer and causes the one or more processor units of the computer to execute the steps of (a) receiving a first data structure comprising said breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCAl*; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE low, LN⁻; (c) receiving a first data structure comprising the level of expression of at least two genes in a cell sample taken from said breast cancer patient wherein markers for said at least two genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER+, ER/AGE high; Table 4 if said patient is classified as ER+, ER/AGE low, LN+; or Table 5 if said patient is classified as ER+, ER/AGE high, LN⁻; (d) determining the similarity of the level of expression of said at least two genes to control levels of expression of said at least two genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity of said level of expression of said genes to said control levels of expression to obtain first and second similarity threshold values, respectively, wherein said second similarity threshold indicates greater similarity to said control levels of expression than does said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said first and said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value or said second threshold similarity value. In a specific embodiment of said computer program product, said first threshold value of similarity and said second threshold value of similarity are values stored in said computer. In another more specific embodiment, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," and wherein said computer program mechanism may be loaded into the memory and further cause said one or more processor units of said computer to execute the step of assigning said breast cancer patient a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile. In another specific embodiment, said computer program mechanism may

be loaded into the memory and further cause said one or more processor units of the computer to execute the steps of receiving a data structure comprising clinical data specific to said breast cancer patient. In a more specific embodiment, said single-channel hybridization intensity values are log transformed. The computer implementation of the method, however, may use any desired transformation method. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the difference between the level of expression of each of said genes in said cell sample taken from said breast cancer patient and the level of expression of the same genes in said control. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the mean log level of expression of each of said genes in said control to obtain a control mean log expression level for each gene, calculating the log expression level for each of said genes in a breast cancer sample from said breast cancer patient to obtain a patient log expression level, and calculating the difference between the patient log expression level and the control mean log expression for each of said genes. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating similarity between the level of expression of each of said genes in said cell sample taken from said breast cancer patient and the level of expression of the same genes in said control, wherein said similarity is expressed as a similarity value. In more specific embodiment, said similarity value is a correlation coefficient. The similarity value may, however, be expressed as any art-known similarity metric.

[00217] Of course, the above breast cancer-specific examples are not limiting; analogous computer systems, software, and data analysis methods may be utilized for any condition of interest. For example, analogous software may be used to determine the prognosis of any other type of cancer, or of any other non-cancer diseases or conditions, using markers, expression level data and controls specific for that cancer, non-cancer disease or condition.

[00218] In an exemplary implementation, to practice the methods of the present invention, a user first loads experimental data into the computer system. These data can be directly entered by the user from a monitor, keyboard, or from other computer systems linked by a network connection, or on removable storage media such as a CD-ROM, floppy disk (not illustrated), tape drive (not illustrated), ZIP[®] drive (not illustrated) or through the network. Next the user causes execution of expression profile analysis software which performs the methods of the present invention.

[00219] In another exemplary implementation, a user first loads experimental data and/or databases into the computer system. This data is loaded into the memory from the storage media or from a remote computer, preferably from a dynamic geneset database system, through the network. Next the user causes execution of software that performs the steps of the present invention.

[00220] Additionally, because the data obtained and analyzed in the software and computer system products of the invention may be confidential, the software and/or computer system preferably comprises access controls or access control routines, such as password protection and preferably, particularly if information is to be transmitted between computers, for example, over the Internet, encryption of the data by a suitable encryption algorithm (*e.g.*, PGP).

[00221] Alternative computer systems and software for implementing the analytic methods of this invention will be apparent to one of skill in the art and are intended to be comprehended within the accompanying claims. In particular, the accompanying claims are intended to include the alternative program structures for implementing the methods of this invention that will be readily apparent to one of skill in the art.

6. EXAMPLE: IDENTIFICATION OF PHENOTYPIC SUBSETS AND INFORMATIVE GENESETS FOR EACH

[00222] Materials and Methods

Tumor Samples:

[00223] 311 cohort samples were collected from breast cancer patients. Selection criteria for sporadic patients (*i.e.*, those not identified as having a *BRCA1*-type tumor; $n = 291$) included: primary invasive breast carcinoma less than 5 cm (T1 or T2); no axillary metastases (N0); age at diagnosis of less than 55 years; calendar year of diagnosis 1983–1996; and no previous malignancies. All patients were treated by modified radical mastectomy or breast-conserving treatment. See van't Veer *et al.*, *Nature* 415:530 (2002). Selection criteria for hereditary (*i.e.*, *BRCA1*-type; $n = 20$) tumors included: carriers of germline mutation in *BRCA1* or *BRCA2*, and primary invasive breast carcinoma. van't Veer, *supra*. Additionally, for development of a classifier for the *BRCA1* group, 14 *BRCA1* samples previously identified (see van't Veer, *supra*) were added to the 20 *BRCA1* type samples to increase sample size. Those 14 samples also satisfy the conditions that they are ER negative and age less than 55 years old.

[00224] Data analysis:

[00225] **Sample sub-grouping:** As shown in FIG. 1, tumor samples were first divided into ER⁺ and ER⁻ branches since this is the dominant gene expression pattern. In the ER⁻ branch, the samples were further divided into “BRCA1 mutation like” and “Sporadic like” categories using the expression templates and 100 genes previously identified as optimal for determining *BRCA1* status. See van’t Veer *et al.*, *Nature* 415:530 (2002). In the ER⁺ category, samples were divided by ER vs. age distribution (see below) into two groups, “ER/AGE low” and “ER/AGE high.” Within the “ER/AGE low” group, samples were further divided according to the lymph node status into two sub-groups: lymph node negative (0 lymph nodes; LN-) and positive (> 0 lymph nodes; LN+) group.

[00226] The result of these divisions was five distinctive sub-groups: “ER⁻, sporadic” ($n = 52$), “ER⁻, BRCA1” ($n = 34$), “ER⁺, ER/AGE high” ($n = 83$), “ER⁺, ER/AGE low, LN⁻” ($n = 81$), and “ER⁺, ER/AGE low, LN⁺” ($n = 75$). A few samples with a specific ER vs. age distribution in “ER⁺, ER/AGE low, LN⁺” group were further excluded to develop a classifier, see below for details.

[00227] **Estrogen receptor level:** Estrogen receptor gene expression level was measured by a 60mer oligo-nucleotide on a microarray. Since every individual sample was compared to a pool of all samples, the ratio to pool was used to measure the relative level. A threshold of -0.65 on $\log_{10}(\text{ratio})$ was used to separate the ER⁺ group from ER⁻ group. See van’t Veer *et al.*, *Nature* 415:530 (2002).

[00228] **Grouping by ER vs. age distribution:** Samples were not uniformly distributed in ER vs. age space among the ER⁺ samples (FIG. 2). First, it appeared that the ER level increases with age, as there were few samples from young individuals having a high ER expression level. For example, in the 35 to 40 years age group, samples having a $\log(\text{ratio})$ of ER > 0.2 are relatively few as compared to the 40 to 45 age group. In the set of samples used, the $40 < \text{age} \leq 45$ group contains 30 samples having $\log(\text{ratio})$ ER values between -0.2 to 0.2 , and 28 samples having values greater than 0.2 , whereas the $35 < \text{age} \leq 40$ group includes 24 samples with values between -0.2 to 0.2 , but only 6 samples with values of greater than 0.2 (Fisher’s exact test P-value: 1%). The increasing ER level with age may simply due to the fact that estrogen levels decrease with age, and the estrogen receptor level rises in compensation.

[00229] There also appeared to be at least two groups of patients, as indicated by the solid line separating the two in FIG. 2A. A bimodality test of the separation indicated by the solid line yielded P-value $< 10^{-4}$. Each of these two groups has its own trend between the ER level and age. The solid line can be approximated by $\text{ER} = 0.1(\text{age} - 42.5)$. Patients having values

above the solid line are referred to as the “ER/AGE high” group, and the patients below the line as the “ER/AGE low” group.

[00230] Prognosis in each group:

[00231] Feature selection and performance evaluation: For the prognosis in each group, non-informative genes were filtered in each group of patients. Specifically, only genes with $|\log_{10}(\text{ratio})| > \log_{10}(2)$ and P-value (for $\log(\text{ratio}) \neq 0$) < 0.01 in more than 3 experiments were kept. This step removed all genes that never had any significant change across all samples. The second step used a leave-one-out cross validation (LOOCV) procedure to optimize the number of reporter genes (features) in the classifier and to estimate the performance of the classifier in each group. The feature selection was included inside the loop of each LOOCV process. The final “optimal” reporter genes were selected using all of the “training samples” as the result of “re-substitution” because one classifier was needed for each group.

[00232] Selection of training samples: Only the samples from patients who had metastases within 5 years of initial diagnosis (3 years for “ER⁻, sporadic” samples; *i.e.*, the “poor outcome” group), or who were metastases-free with more than 5 years of follow-up time (*i.e.*, the “good outcome” group), were used as the training set. Because the average expression levels for informative genes among patients who were metastasis-free, or who had early metastases, were used as expression templates for prediction, the training samples for the ER⁺ samples were further limited to those samples that could also be correctly classified by the first round of LOOCV process. For the “ER⁻, sporadic” samples, no such iteration was done because no improvement was observed. For the “ER⁻, BRCA1” samples, an iteration was done, but the training samples in the second iteration were limited to the correctly predicted good outcome samples from the first round of LOOCV, and all the poor outcome samples with metastases time less than 5 years. Further limitation of the poor outcome samples was not performed because of the small number of poor samples and the absence of improvement by such limitation. In the first round of LOOCV, except for the “ER⁻, sporadic” group, the number of features was fixed at 50 genes. A patient was predicted to have a favorable outcome, that is, no metastases within five years of initial diagnosis, if the expression of the reporter genes in a sample from the individual was more similar to the “average good profile” than the “average poor profile”, and a poor outcome, that is, a metastasis within five years, if the expression of the reporter genes in the sample was more similar to the “average poor profile” than the “average good profile”.

[00233] The justification for such an iteration operation is threefold. First, biologically, there are always a few individuals with specific reasons (different from the vast majority) to stay metastases free or to develop metastases. Second, statistically, most groups of patients include outliers that don't follow the distribution of the majority of samples. Third, methodologically, the iteration operation is very similar to the idea of "boosting", but instead of increasing the weights of the samples predicted wrong, emphasis is placed on the well behaved samples for selecting features and training the classifier. Since this process was used to select "training samples", and the performance was evaluated using the LOOCV (including the feature selection) after the training sample being fixed, there is no issue of over-fitting involved in our procedures. This method of iteration is thus more likely to reveal the dominant mode to metastases within each group.

[00234] Error rate and odds ratio, threshold in the final LOOCV: Unless otherwise stated, the error rate was the average error rate from two populations: (1) the number of poor outcome samples misclassified as good outcome samples, divided by the total number of poor outcome samples; and (2) the total number of good outcome samples misclassified as poor outcome samples, divided by the total number of good samples. Two odds ratios were reported for a given threshold: (1) the overall odds ratio and (2) the 5 year odds ratio. The 5 year odds ratio was calculated from samples from individuals that were metastases free for more than five years, and who experienced metastasis within 5 years. The threshold was applied to $\text{cor1} - \text{cor2}$, where "cor1" stands for the correlation to the "average good profile" in the training set, and "cor2" stands for the correlation to the "average poor profile" in the training set.

[00235] The threshold in the final round of LOOCV was defined using the following steps: (1) For each of the N sample i left out for training, features based on the training set were selected, (2) given a feature set, an incomplete LOOCV with $N-1$ samples was performed (only the "average poor profile" and "average good profile" is varied depending on whether the left out sample is in the training set or not), (3) the threshold based on the minimum error rate from $N-1$ samples was determined, and that threshold was assigned to sample i in step (1), (4) the median threshold from all N samples was taken, and designated the final threshold. FIGS. 3-7 present detailed information about classifiers for the 5 groups: "ER⁻, sporadic", "ER⁻, BRCA1", "ER⁺, ER/age high", "ER⁺, ER/age low, LN⁻", "ER⁺, ER/age low, LN⁺". Tables 1-5 (see Section 5.3) list the final optimal reporter genes for each of the 5 classifiers for each of the five patient subsets. Table 6, below, summarizes the performance of each of the five classifiers together with thresholds used in each classifier.

[00236] Table 6. Performance of classifiers for each patient subset.

Classifier	Optimal # of Genes	(C1-C2) Threshold	Metastasis Free	# of Samples	TP	FP	FN	TN	Odds Ratio	95% C.I.
ER+, ER/AGE high	50	1.22	Overall	83	31	14	5	33	14.61	4.71-45.36
			5 year	71	24	11	3	33	24.00	6.03-95.46
ER+, ER/AGE low, LN-	65	0.38	Overall	81	14	6	6	55	21.39	5.98-76.52
			5 year	73	11	4	5	53	29.15	6.73-126.33
ER+, ER/AGE low, LN+	50	-0.12	Overall	56	7	4	6	39	11.38	2.54-50.94
			5 year	48	5	4	3	36	15.00	2.57-87.64
ER-, sporadic	20	-0.01	Overall	52	18	7	7	29	7.35	2.16-25.04
			5 year	45	16	5	6	18	9.60	2.45-37.58
ER-, BRCA1	10	-0.37	Overall	34	6	3	3	22	14.67	2.34-92.11
			5 year	22	6	1	3	12	24.00	2.04-282.68

[00237] TP: True positive

[00238] FP: False positive

[00239] FN: False negative

[00240] TN: True negative

[00241] Classification method: All classifiers described herein, feature selection and optimization were included inside the LOOCV loop. Classifier performance was based on the LOOCV results. The profile based on the selected features from each patient was compared to the “average good profile” and “average poor profile” (by correlation) to determine its predicted outcome.

[00242] Correlation calculation: The correlation between each gene's expression $\log(\text{ratio})$ and the endpoint data (final outcome) was calculated using the Pearson's correlation coefficient. The correlation between each patient's profile and the “average good profile” and “average poor profile” was the cosine product (no mean subtraction).

[00243] Results:

[00244] The comprehensive prognosis strategy was employed on microarray expression profiles of 311 patients diagnosed before age 55 that were all part of previous studies establishing and validating a 70-gene prognosis profile. See van 't Veer *et al.*, *Nature* 415:530 (2002); van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002). In addition, 14 known *BRCA1* samples from the *Nature* study were included in defining the prognosis

classifier for the *BRCA1* group. The overview of the stratifications is shown in FIG. 1. In each of the patient subsets, prognosis classifiers were developed and performance was evaluated by leave-one-out cross-validation. The biological make up of each of the classifiers was also examined.

[00245] During the process to decide whether a particular clinical parameter should be used for the next stratification, our objectives were twofold: (1) identification of homogeneous prognosis patterns; and/or (2) improved prognosis in the subsets. There is a subtle balance between these two objectives because smaller groups will likely lead to uniform patterns within the group but have increasingly limited predictive power. With the exception of the *BRCA1* subset, each group in our stratification contained 50 or more samples.

[00246] The first layer of stratification was based on the estrogen receptor level. It was previously observed that estrogen receptor expression has a dominant effect on overall gene expression in breast cancer as seen in hierarchical clustering. van 't Veer *et al.*, *Nature* 415:530 (2002); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001). In previous analysis up to 2500 genes were significantly correlated with ER expression levels in tumor. See, van 't Veer *et al.*, *Nature* 415:530 (2002). According to the threshold defined previously (van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002)), samples were first divided into two groups according to the estrogen receptor level as measured by the oligo probe (accession number: NM_000125) on the array; samples with $\log(\text{ratio}) > -0.65$ belong to the ER⁺ group, and the rest belong to ER⁻ group). This resulted in 239 samples in the ER⁺ group and 72 samples in the ER⁻ group.

[00247] In the ER⁺ branch it was observed that when displaying ER expression level as a function of age, at least two subgroups appeared to exist. (In general, any bimodality in the clinical data is useful.) The tumors were stratified according this bimodality (see FIG. 2). The group of ER⁺ patients having a high ER/AGE ratio was designated the “ER/AGE high” group (83 samples), and the remaining group of patients was designated “ER/AGE low” group (156 samples).

[00248] Within the “ER/age high” group, a group of prognosis reporter genes that highly correlated with the outcome is identified (see Table 3). Moreover, the expression of these genes appeared to be very homogeneous, as indicated by high similarity in expression among those genes. See FIG 2A. Leave-one-out cross validation including reporter selection yielded an odds ratio of 14.6 (95%CI: 4.7-45.4) and 5 year odds ratio of 24.0 (95%CI: 6.0-95.5). Examination of those reporter genes reveals they are mostly the cell cycle genes which are highly expressed in the poor outcome tumors. It is worth noting that even though this

group includes LN+ and LN- individuals, and mixed treatment, the incidence of distant metastases is predicted by a biologically uniform set of genes, possibly indicating that proliferation is the prime driving force for disease progression. Also even though variation in these genes is observed in other tumor subgroups this is generally not correlated with outcome in those settings (see below).

[00249] In the “ER/age low” group, no predictive pattern was found in the whole group; thus, the samples were further stratified into LN- (81 samples, referred to as “ER/age low LN-”) and LN+ (75 samples, referred to as “ER/age low LN+”) group.

[00250] Within the “ER/age low LN-” group, a group of genes was identified that was uniformly co-regulated, and which correlated with the outcome. Leave-one-out cross-validation (including feature selection) yielded an odds ratio of 21.4 (95% CI: 6.0-76.5) and 5 year odds ratio of 29.2 (95% CI: 6.7-126.3). This group of genes is also enriched for individual biological functions (see below).

[00251] For the “ER/age low LN+” subset, an informative set of genes (*see* Table 4) was obtained after exclusion of several samples from older individuals having low ER levels. These samples are indicated in FIG. 2A as those lying below the dashed line (approximated as $ER < 0.1 * (age - 50)$). 56 samples remained after the exclusion. This sample set allowed the identification of a group of genes with a highly homogeneous pattern that is useful for prognosis (overall odds ratio: 11.4 (2.5-50.9), 5 year odds ratio: 15.0 (2.6-87.6)). This suggests again that ER vs. age is an important combination for stratifying breast cancer patients. The reporter genes involved in this classifier also correlated with the clinical measure of the degree of lymphocytic infiltration (data not shown). The prediction in this group was not as strong as other positive groups, which may indicate the primary tumor carries weaker information about the metastases for this group of patients, and the metastases may be started from or influenced by tumors already in lymph nodes.

[00252] In the ER- branch, because a portion of the samples are “BRCA1-like,” it is natural to divide the samples into “BRCA1-like” and “sporadic like”. To perform the classification, the BRCA1/sporadic tumor type classifier described in Roberts *et al.*, “Diagnosis and Prognosis of Breast Cancer Patients,” International Publication No. WO 02/103320, which is hereby incorporated by reference in its entirety, to segregate the ER- cohort samples. 52 out of the 72 ER- samples were found to be “sporadic like” and 20 were found to be “BRCA1-like”. Interestingly, the “sporadic like” group was enriched for erbb2 mutations (data not shown).

[00253] Within the “ER⁻, sporadic” group, no homogeneous prognosis pattern was identified; however, 20 genes were identified that are highly predictive of the tumor outcome (see Table 2). Leave-one-out cross-validation including feature selection yielded an odds ratio of 7.4 (95% CI 2.2-25.0) and 5 year odds ratio 9.6 (2.5 – 37.6). This result represents a significant improvement in prognosis compared to the previously-identified 70 gene prognosis classifier (see Roberts *et al.*, International Publication No. WO 02/103320; van't Veer *et al.*, *Nature* 415:530 (2002)) which has no within-group prognostic power for the ER⁻ patient subset. The fact that 20 genes predict outcome and that there is no homogeneous (and apparent biological) pattern in this group probably indicates multiple mechanisms of metastasis in this group. Gene annotation indicates that genes included may be involved in invasion, energy metabolism and other functions.

[00254] For the “ER⁻, BRCA1-like” group, we added 14 BRCA1 mutation carrier samples from a previous study were added to increase the number of samples. Those 14 extra samples also satisfied the following selection criteria: ER negative and age less than 55 years. The leave-one-out cross validation process identified 10 genes that are predictive of final outcomes. The overall odds ratio is 14.7 (95% CI: 2.3-92.1) and the 5 year odds ratio is 24.0 (95% CI: 2.0-282.7).

[00255] Because no homogeneous gene expression patterns were found in ER⁻ branch, the predictive power of those genes was further validated. One means of further validation was to review the different classifier gene sets for biological interpretations and to identify genes within each classifier that gave indications as to the origins of the tumors.

[00256] The “ER⁺, ER/AGE high” group yielded a classifier highly enriched for cell cycle genes with both G1/S and G2/M phases represented. In this group, over-expression of 46 of the 50 genes was associated with disease progression including all the known cell cycle genes. This is consistent with rapid growth being the determinant of metastatic potential. Four genes in this classifier were anti-correlated with outcome and cell cycle. One of these genes encodes follistatin, which binds to and inhibits activin and other members of the TGFβ family (Lin *et al.*, *Reproduction* 126:133 (2003)), the members of which have many functions, including growth stimulation. Tumor grade also accurately predicted metastatic potential in this group (overall odds ratio: 5.9, 95% CI: 2.0-18.0, 5 year odds ratio: 12.5, 95% CI: 2.6-59.3) and was also correlated with the expression level of these genes, which is consistent with rate of growth being the primary determinant of disease progression. This set of genes had a significantly lower correlation with outcome in the other patient subsets, even though coordinate and similarly variable expression was seen. For example, many tumors in

the “ER⁻, sporadic” group had high cell cycle and low FST expression, but the expression of these genes in these groups was minimally correlated with outcome, indicating that growth was not the primary determinant of outcome here (*see* FIGS. 8A and 8B).

[00257] The ER⁺, ER/AGE low, LN⁻ group yielded a classifier rich in both genes for glycolytic enzymes (12 of 56) and genes induced by hypoxia and/or angiogenesis (14 of 56) with 5 genes falling into both categories. These genes were positively correlated with poor outcome, implying that energy metabolism (glycolysis), angiogenesis and adaptation to hypoxia were critical pathways in this subgroup of tumors. None of these genes appeared in the classifiers for the other patient subsets, and there was a much reduced predictive value of these genes in the other tumors, even though coordinate and similarly variable expression was seen (*see* FIG. 8C and 8D).

[00258] The implication of the above analyses is that certain well known functions (growth, angiogenesis, energy metabolism) are important in certain tumor types and not in others, and therefore therapies that target these functions will be likely be similarly effective in some tumor subgroups and not in others. For example therapies that target cell cycle progression, such as taxol or the vinca alkaloids, may be optimally effective in the ER⁺, ER/AGE high group, where overexpression of cell cycle genes predominates in the classifier. In contrast, tumor subgroups in which variation in cell cycle expression is not correlated with outcome may be less sensitive to taxol or the vinca alkaloids.

[00259] The “comprehensive prognosis” approach significantly improved the prediction error rate when compared with 70 gene classifier (Table 7). To make the comparison fair, we listed two sets of results from the 70 gene classifier. The first results from the use of the same threshold applied to all the patient subsets (threshold previously optimized for false negative rate); the second one results from the use of a threshold optimized for each patient subset (optimized for average error rate). The comprehensive approach lowered the error rate by at least 6%.

Table 7. Average error rate for the patient subset approach compared with the previously-described 70 gene classifier.

Prognosis method	over all error rate	5 year error rate
70 gene, fix thresh	30.90%	25.70%
70 gene, opt thresh	28.60%	27.60%
Comprehensive	21.50%	19.30%

[00260] Fix thresh: use of a fixed threshold in the classifier as previously determined.

[00261] Opt threshold: use of a threshold optimized for each sub-group. For the “ER/Age low, LN+” subgroup, 56 samples used for developing the classifier were included here, resulted in 306 samples in total.

[00262]

7. REFERENCES CITED

[00263] All references cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual publication or patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

[00264] Many modifications and variations of the present invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be limited only by the terms of the appended claims along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics of said condition into a plurality of first classes; and

(b) identifying within each of said first classes a first set of genes or markers informative for said condition

wherein said first set of genes or markers within each of said first classes is unique to said class relative to other first classes.

2. The method of claim 1, which further comprises additionally classifying into a plurality of second classes said samples or individuals in at least one of said first classes on the basis of a phenotypic or genotypic characteristic different than that used in said classifying step (a); and identifying within at least one of said second classes a second set of informative genes or markers, wherein said second set of informative genes or markers within each of said second classes is unique to said second class relative to other first and second classes.

3. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) classifying each of a plurality of samples or individuals on the basis of one or more phenotypic or genotypic characteristics into a plurality of first classes;

(b) classifying at least one of said first classes into a plurality of second classes on the basis of phenotypic or genotypic characteristic different than that used in said classifying step (a); and

(c) identifying within at least one of said first classes or said second classes a set of genes or markers informative for said condition,

wherein said second set of genes or markers is unique to said class relative to other first and second classes.

4. A method of identifying a set of informative genes or markers for a condition comprising a plurality of phenotypic or genotypic characteristics, comprising:

(a) selecting a first characteristic from said plurality of phenotypic or genotypic characteristics;

(b) identifying at least two first condition classes differentiable by said first characteristic;

(c) selecting a plurality of individuals classifiable into at least one of said first condition classes; and

(d) identifying in samples derived from each of said plurality of individuals a set of genes or markers informative for said condition within said at least one of said first condition classes.

5. A method of classifying an individual with a condition as having a good prognosis or a poor prognosis, comprising:

(a) classifying said individual into one of a plurality of patient classes, said patient classes being differentiated by one or more phenotypic, genotypic or clinical characteristics of said condition;

(b) determining the level of expression of a plurality of genes or their encoded proteins in a cell sample taken from the individual relative to a control, said plurality of genes or their encoded proteins comprising genes or their encoded proteins informative for prognosis of the patient class into which said individual is classified; and

(c) classifying said individual as having a good prognosis or a poor prognosis on the basis of said level of expression.

6. The method of claim 5, wherein said condition is cancer, said good prognosis is the non-occurrence of metastases within five years of initial diagnosis, and said poor prognosis is the occurrence of metastases within five years of initial diagnosis.

7. The method of claim 5, wherein said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a poor prognosis.

8. The method of claim 7, in which said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a poor prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance.

9. The method of claim 5, wherein said control is the average level of expression of each of said plurality of genes or their encoded proteins across a plurality of samples derived from individuals identified as having a good prognosis.

10. The method of claim 9, in which said classifying in step (c) is carried out by a method comprising comparing the level expression of each of said plurality of genes or their encoded proteins to said average level of expression of each corresponding gene or its encoded protein in said control, and classifying said individual as having a good prognosis if said level of expression correlates with said average level of expression of each of said genes or their encoded proteins in said control more strongly than would be expected by chance.

11. The method of claim 5, wherein said plurality of patient classes comprises ER^{-} , *BRCA1* individuals; ER^{-} , sporadic individuals; ER^{+} , ER/AGE high individuals; ER^{+} , ER/AGE low, LN^{+} individuals; and ER^{+} , ER/AGE low, LN^{-} individuals.

12. A method of classifying a breast cancer patient as having a good prognosis or a poor prognosis comprising:

(a) classifying said breast cancer patient as ER^{-} , *BRCA1*; ER^{-} , sporadic; ER^{+} , ER/AGE high; ER^{+} , ER/AGE low, LN^{+} ; or ER^{+} , ER/AGE low, LN^{-} ;

(b) determining the level of expression of a first plurality of genes in a cell sample taken from said breast cancer patient relative to a control, said first plurality of genes comprising two of the genes corresponding to the markers in Table 1 if said breast cancer patient is classified as ER^{-} , *BRCA1*; in Table 2 if said breast cancer

patient is classified as ER⁻ sporadic; in Table 3 if said breast cancer patient is classified as ER⁺, ER/AGE high; in Table 4 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁺; or in Table 5 if said breast cancer patient is classified as ER⁺, ER/AGE low, LN⁻; and

(c) classifying said breast cancer patient as having a good prognosis or a poor prognosis on the basis of the level of expression of said first plurality of genes,

wherein said breast cancer patient is “ER/AGE high” if the ratio of the $\log_{10}(\text{ratio})$ of ER gene expression to age exceeds a predetermined value, and “ER/AGE low” if the ratio of the $\log_{10}(\text{ratio})$ of ER gene expression to age does not exceed said predetermined value.

13. The method of claim 12, wherein said control is the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, *BRCA1* individuals, if said breast cancer patient is ER⁻, *BRCA1*; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁻, sporadic individuals if said breast cancer patient is ER⁻, sporadic; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE high individuals, if said breast cancer patient is ER⁺, ER/AGE high; the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁺ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁺; or the average level of expression of each of said plurality of genes in a plurality of samples derived from ER⁺, ER/AGE low, LN⁻ individuals where said breast cancer patient is ER⁺, ER/AGE low, LN⁻.

14. The method of claim 13, wherein each of said individuals has a poor prognosis.

15. The method of claim 13, wherein each of said individuals has a good prognosis.

16. The method of claim 14, wherein said classifying step (c) is carried out by a method comprising comparing the level of expression of each of said plurality of genes or their encoded proteins in a sample from said breast cancer patient to said control, and classifying said breast cancer patient as having a poor prognosis if said level of expression

correlates with said average level of expression of the corresponding genes or their encoded proteins in said control more strongly than would be expected by chance.

17. The method of claim 12, wherein said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said individual.

18. The method of claim 12, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises two of the genes for which markers are listed in Table 1.

19. The method of claim 12, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

20. The method of claim 12, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises two of the genes for which markers are listed in Table 2.

21. The method of claim 12, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

22. The method of claim 12, wherein said individual is ER⁺, ER/AGE high, and said plurality of genes comprises two of the genes for which markers are listed in Table 3.

23. The method of claim 12, wherein said individual is ER⁺, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

24. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises two of the genes for which markers are listed in Table 4.

25. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

26. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises two of the genes for which markers are listed in Table 4.

27. The method of claim 12, wherein said individual is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

28. The method of claim 12, further comprising determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

29. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising:

(a) classifying said individual as ER⁻, *BRCAl*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻;

(b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCAl*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻;

(c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and

(d) assigning said individual to one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis.

30. The method of claim 29, wherein said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a).

31. The method of claim 29, wherein said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer.

32. The method of claim 29, further comprising determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of

genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

33. A method of identifying a set of genes informative for a condition, said condition having a plurality of phenotypic or genotypic characteristics such that samples may be categorized by at least one of said phenotypic or genotypic characteristics into at least one characteristic class, said method comprising:

- (a) selecting a plurality of samples from individuals having said condition;
- (b) identifying a first set of genes informative for said characteristic class using said plurality of samples;
- (c) predicting the characteristic class of each of said plurality of samples;
- (d) discarding samples for which said characteristic class is incorrectly predicted;
- (e) repeating steps (c) and (d) at least once; and
- (f) identifying a second set of genes informative for said characteristic class using samples in said plurality of samples remaining after step (e).

34. The method of claim 6, wherein said cancer is breast cancer.

35. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising:

- (a) classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition;
- (b) determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category;
- (c) determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and
- (d) assigning the individual to a category in a clinical trial on the basis of prognosis.

36. A method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising:

(a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class; and

(b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class,

wherein, if a set of genes or markers informative for said condition in individuals in said class is obtained for any of said one or more of said plurality of classes, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

37. The method of claim 36, further comprising, for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of individuals or samples derived therefrom in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein, for each of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for said condition in said organism.

38. A method for identifying one or more sets of informative genes or markers for a condition in an organism, comprising:

(a) subdividing a plurality of individuals or samples derived therefrom of said organism subject to said condition into a plurality of classes based on one or more clinical, phenotypic or genotypic characteristics of said organism, wherein each said class consists of a plurality of individuals or samples derived therefrom of said organism each having said one or more clinical, phenotypic or genotypic characteristics specific for said class;

(b) attempting to identify for each of one or more of said plurality of classes a set of genes or markers informative for said condition in individuals in said class, wherein if a set of genes or markers informative for said condition in individuals in said class is identified for any of said one or more of said classes, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism; and

(c) for each of one or more of said classes in which a set of genes or markers informative for said condition in individuals in said class cannot be obtained, repeating said steps (a) and (b) on said plurality of individuals or samples derived therefrom in said class such that said plurality of samples or individuals in said class is subdivided into a plurality of additional classes based on one or more clinical, phenotypic or genotypic characteristics of said organism which are different from those used for defining said class, wherein, for each of one or more of said plurality of additional classes, if a set of genes or markers informative for said condition in individuals in said class is obtained, said set of genes or markers is taken as a set of informative genes or markers for a condition in said organism.

39. The method of claim 38, wherein said condition is a type of cancer, and wherein each of said sets of genes or markers is informative of prognosis of individuals in a corresponding class.

40. The method of claim 39, wherein said condition is breast cancer, and wherein said one or more clinical, phenotypic or genotypic characteristics comprises age, ER level, ER/AGE, BRAC1 status, and lymph node status.

41. The method of claim 39, further comprising generating a template profile comprising measurements of levels of genes or markers of said set for said class representative of levels of the genes or markers in a plurality of patients having a chosen prognosis level.

42. A method for predicting a breast cancer patient as having a good prognosis or a poor prognosis, comprising:

(a) classifying said breast cancer patient into one of the following classes: (a1) ER⁻, *BRCA1*; (a2) ER⁻, sporadic; (a3) ER⁺, ER/AGE high; (a4) ER⁺, ER/AGE low, LN⁺; or (a5) ER⁺, ER/AGE low, LN⁻;

(b) determining a profile comprising measurements of a plurality of genes or markers in a cell sample taken from said breast cancer patient, said plurality of genes markers comprising at least two of the genes or markers corresponding to the markers in (b1) Table 1 if said breast cancer patient is classified as ER^- , *BRCA1*; (b2) Table 2 if said breast cancer patient is classified as ER^- sporadic; (b3) Table 3 if said breast cancer patient is classified as ER^+ , ER/AGE high; (b4) Table 4 if said breast cancer patient is classified as ER^+ , ER/AGE low, LN^+ ; or (b5) Table 5 if said breast cancer patient is classified as ER^+ , ER/AGE low, LN^- ; and

(c) classifying said breast cancer patient as having a good prognosis or a poor prognosis based on said profile of said plurality of genes or markers,

wherein ER^+ designates a high ER level and ER^- designates a low ER level, wherein said ER/AGE is a metric of said ER level relative to the age of said patient, and wherein LN^+ designates a greater than 0 lymph nodes status in said patient and LN^- designates a 0 lymph nodes status in said patient.

43. The method of claim 42, wherein step (c) is carried out by a method comprising comparing said profile to a good prognosis template and/or a poor prognosis template, and wherein said patient is classified as having a good prognosis if said profile has a high similarity to a good prognosis template or has a low similarity to a poor prognosis template or as having a poor prognosis if said profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of genes or markers representative of levels of said genes or markers in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-reoccurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a patient who has reoccurrence of metastases within a second period of time after initial diagnosis.

44. The method of claim 43, further comprising determining said profile, said ER level, said LN status, and/or, said ER/AGE .

45. The method of claim 44, wherein said profile is an expression profile comprising measurements of a plurality of transcripts in a sample derived from said patient, wherein said

good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

46. The method of claim 45, wherein said expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

47. The method of claim 43, wherein said profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

48. The method of claim 46, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

49. The method of claim 48, wherein similarity of said expression profile to said good prognosis template is represented by a correlation coefficient between said expression profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

50. The method of claim 48, wherein similarity of said expression profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

51. The method of claim 49, wherein said correlation threshold is 0.5.

52. The method of claim 51, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

53. The method of claim 52, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

54. The method of claim 53, wherein said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

55. The method of claim 54, wherein said estrogen receptor level is measured by a polynucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

56. The method of claim 55, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

57. The method of claim 55, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

58. The method of claim 42, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 1.

59. The method of claim 42, wherein said individual is ER⁻, *BRCA1*, and said plurality of genes comprises all of the genes for which markers are listed in Table 1.

60. The method of claim 42, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 2.

61. The method of claim 42, wherein said individual is ER⁻, sporadic, and said plurality of genes comprises all of the genes for which markers are listed in Table 2.

62. The method of claim 42, wherein said individual is ER+, ER/AGE high, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 3.

63. The method of claim 42, wherein said individual is ER+, ER/AGE high, and said plurality of genes comprises all of the genes for which markers are listed in Table 3.

64. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4.

65. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN+, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

66. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises at least two of the genes for which markers are listed in Table 4.

67. The method of claim 42, wherein said individual is ER+, ER/AGE low, LN⁻, and said plurality of genes comprises all of the genes for which markers are listed in Table 4.

68. The method of claim 42, wherein said profile further comprises one or more genes for which markers are not found in Tables 1-5, wherein said one or more genes are informative for prognosis.

69. A method for assigning an individual to one of a plurality of categories in a clinical trial, comprising assigning said individual to one category in a clinical trial if said individual has a good prognosis as determined by the method of any one of claims 7-33, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis as determined by the method of any one of claims 7-33.

70. The method of claim 69, wherein said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a).

71. The method of claim 69, wherein said individual is additionally assigned to a category in said clinical trial on the basis of one or more other clinical, phenotypic or genotypic characteristic of breast cancer.

72. The method of claim 69, further comprising determining in said cell sample the levels of expression of said one or more genes for which markers are not found in Tables 1-5, and determining from said expression levels of said one or more genes, whether said individual has a good prognosis or a poor prognosis.

73. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in any one of Tables 1-5.

74. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 1.

75. The microarray of claim 74, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to each of the genes listed in Table 1.

76. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 2.

77. The microarray of claim 76, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 2.

78. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 3.

79. The microarray of claim 78, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 3.

80. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 4.

81. The microarray of claim 80, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 4.

82. A microarray comprising a plurality of polynucleotide probes each complementary and hybridizable to a sequence in a different gene listed in Table 5.

83. The microarray of claim 82, wherein said plurality of polynucleotide probes comprises a probe complementary and hybridizable to a sequence in each of the genes listed in Table 5.

84. The microarray of any of claims 73-83, wherein said plurality of polynucleotide probes constitutes at least 50% of the probes on said microarray.

85. The microarray of any of claims 73-83, wherein said plurality of polynucleotide probes constitutes at least 90% of the probes on said microarray.

86. The microarray of claim 73, wherein said plurality of polynucleotide probes comprises probes complementary and hybridizable to at least 75% of the genes listed in Table 1, Table 2, Table 3, Table 4, or Table 5, wherein said plurality of polynucleotide probes, in total, constitutes at least 50% of the probes on said microarray.

87. A kit comprising the microarray of any one of claims 73-83 in a sealed container.

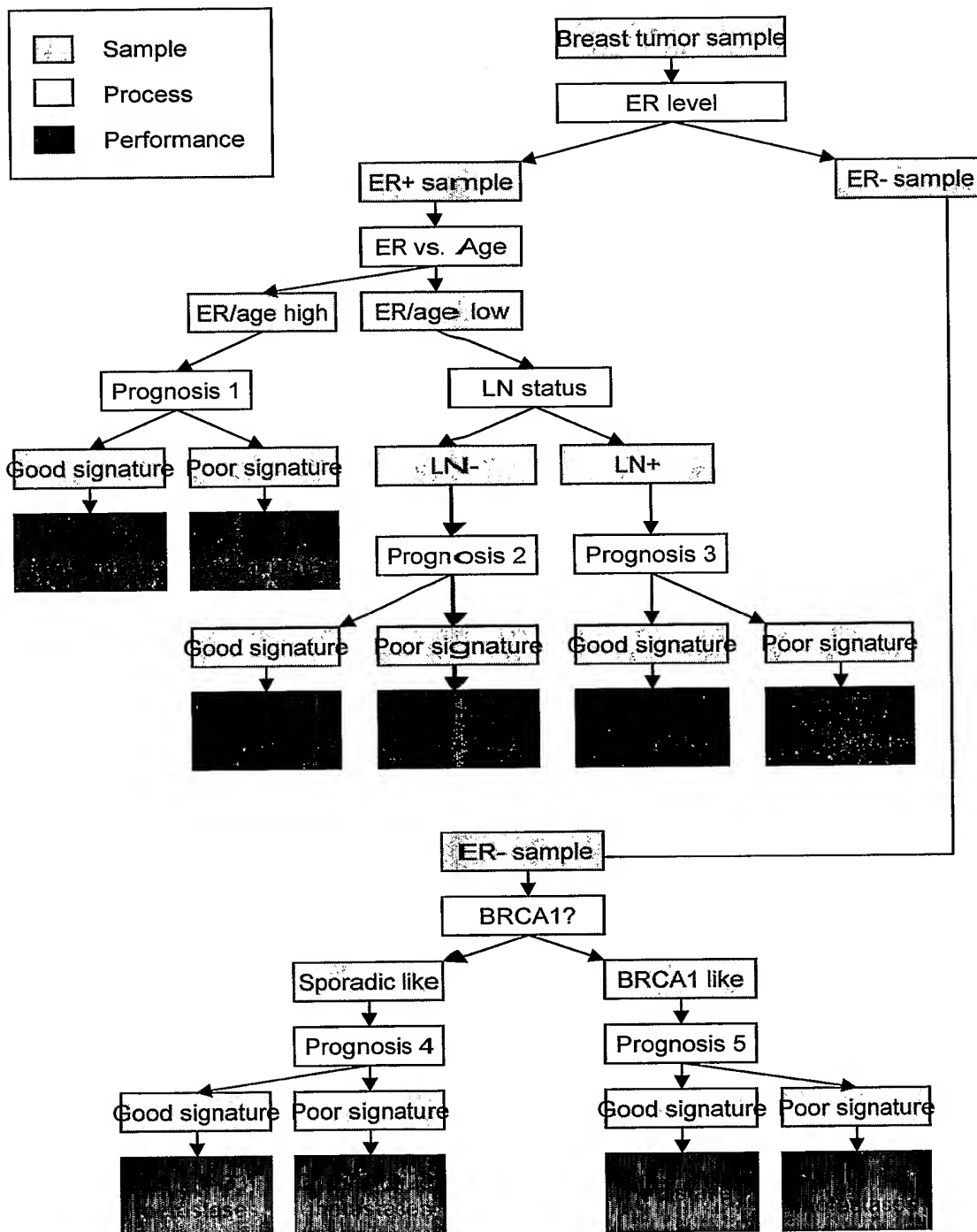


FIG. 1

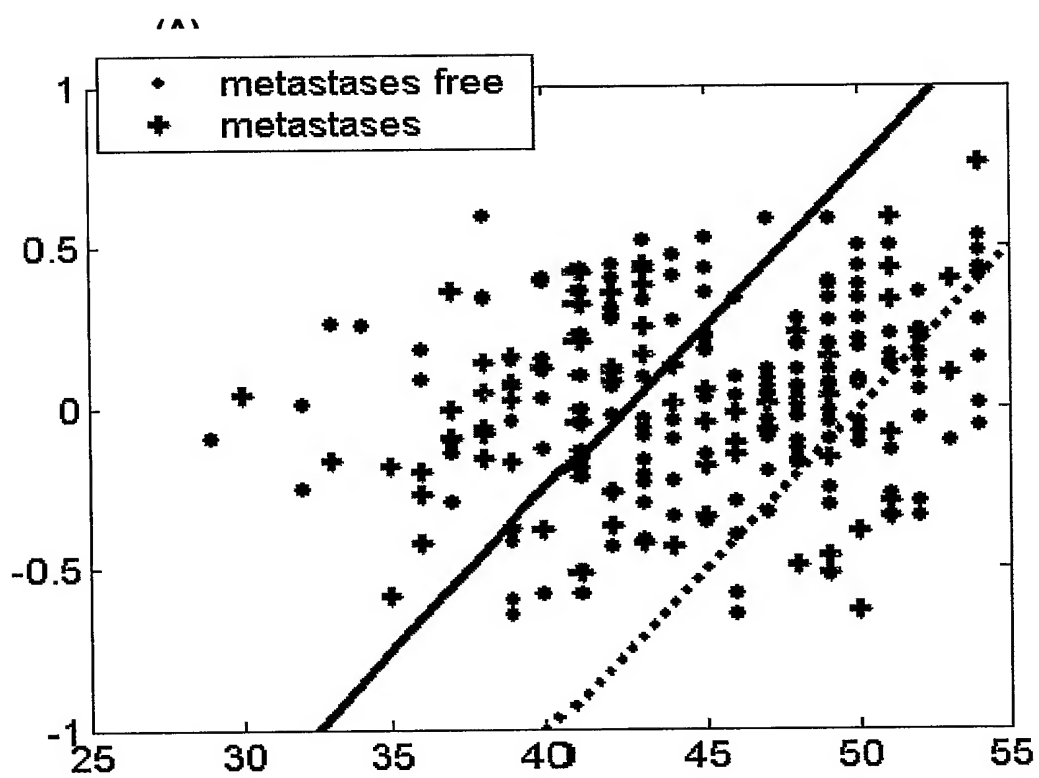
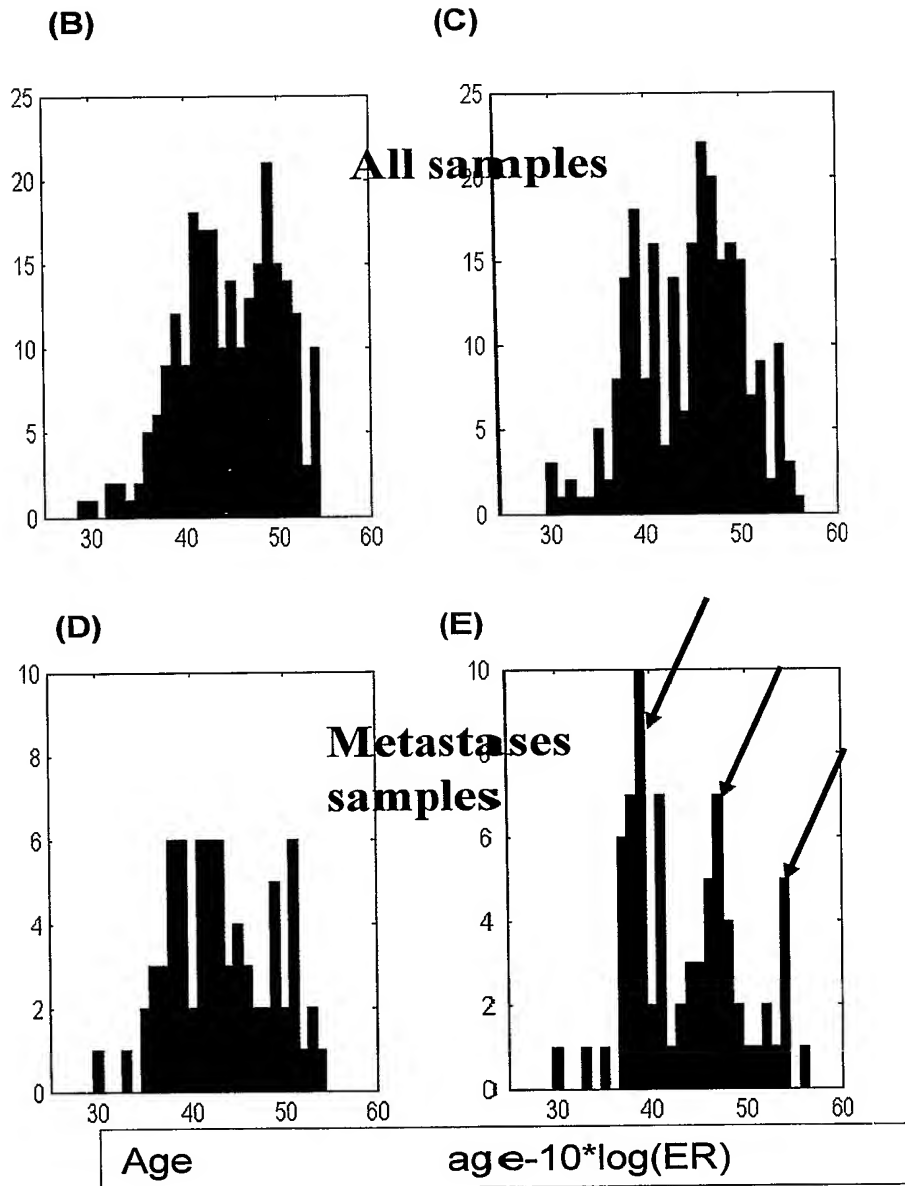
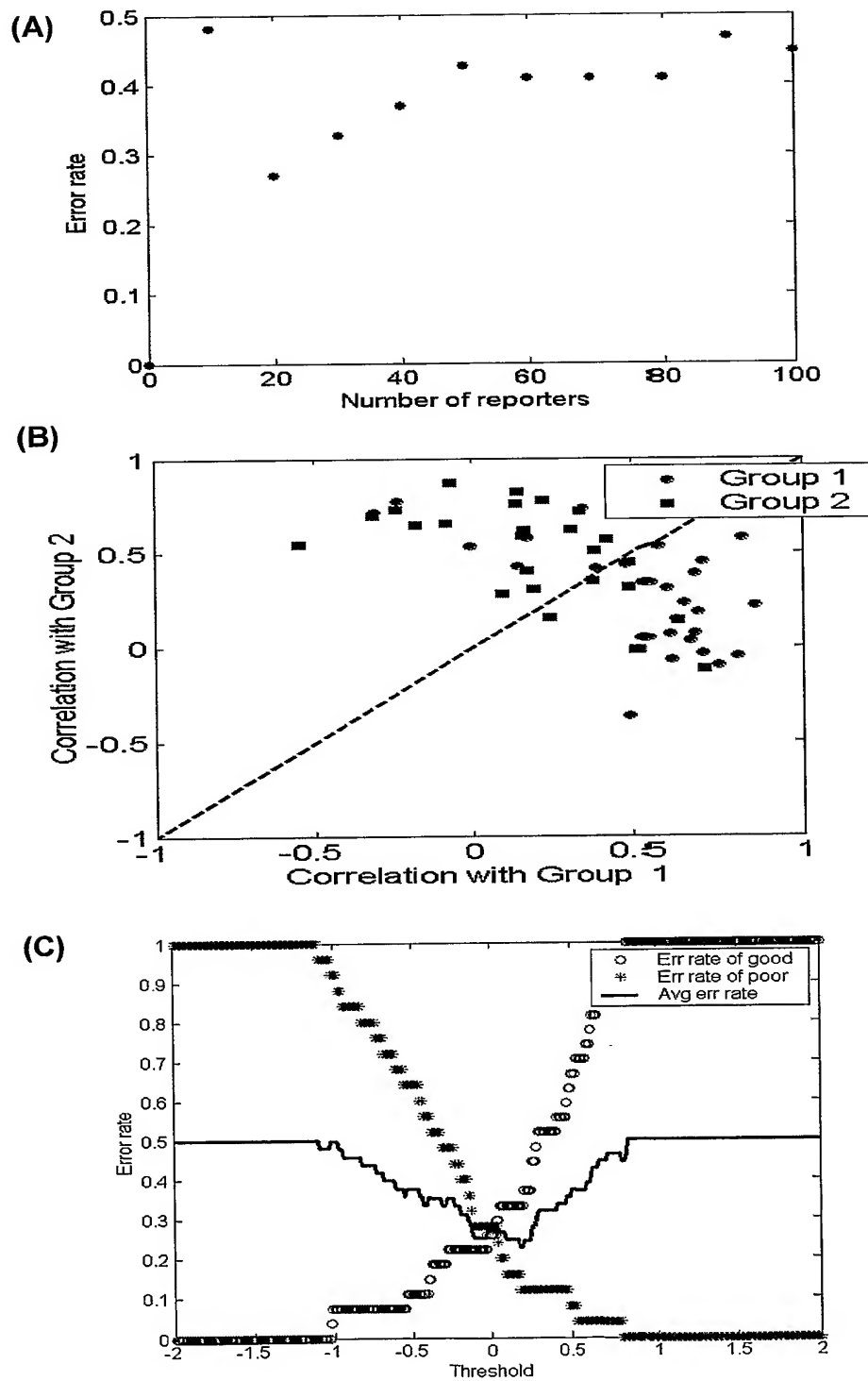


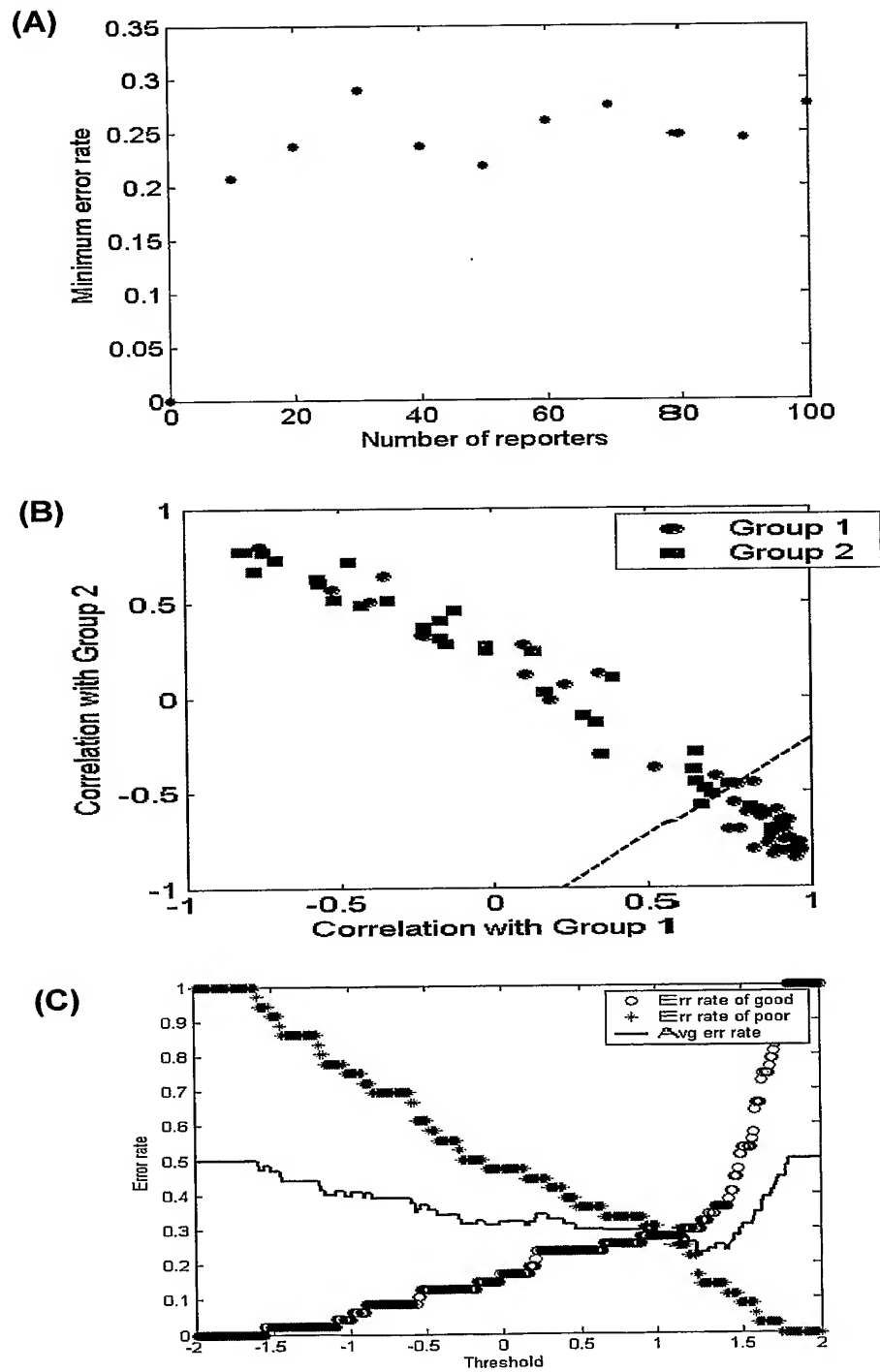
FIG. 2A



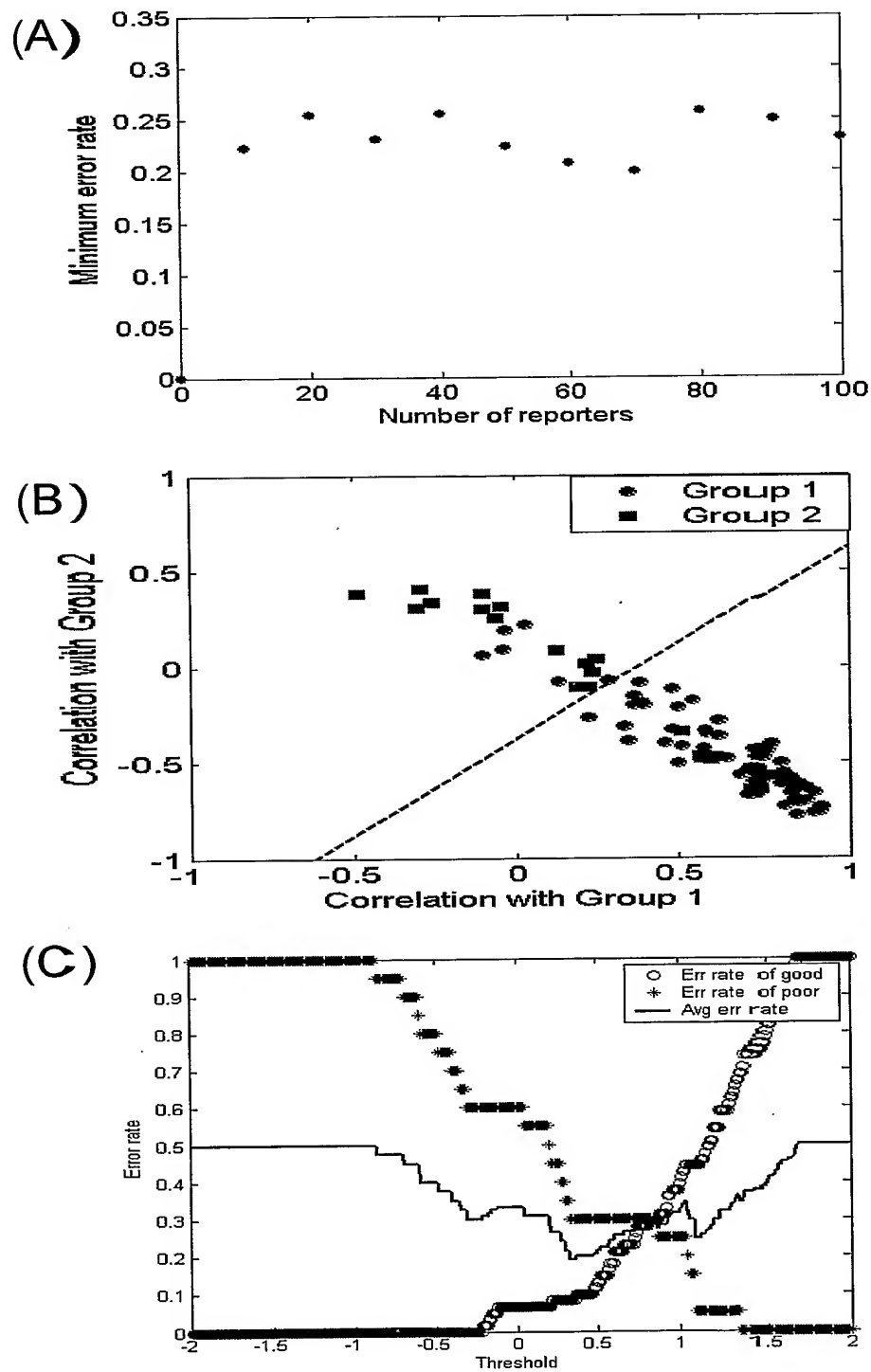
FIGS. 2B-E



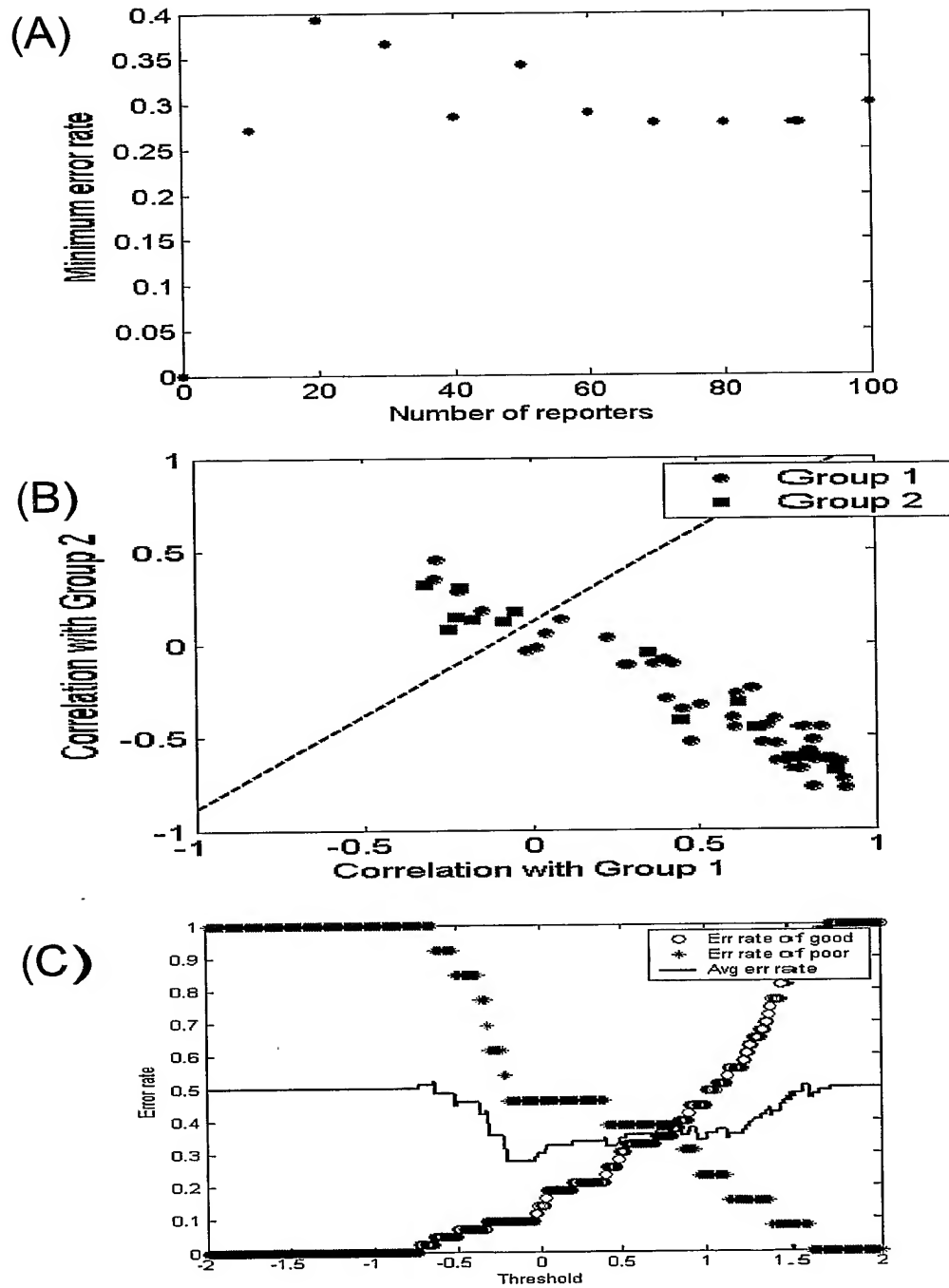
FIGS. 3A-C



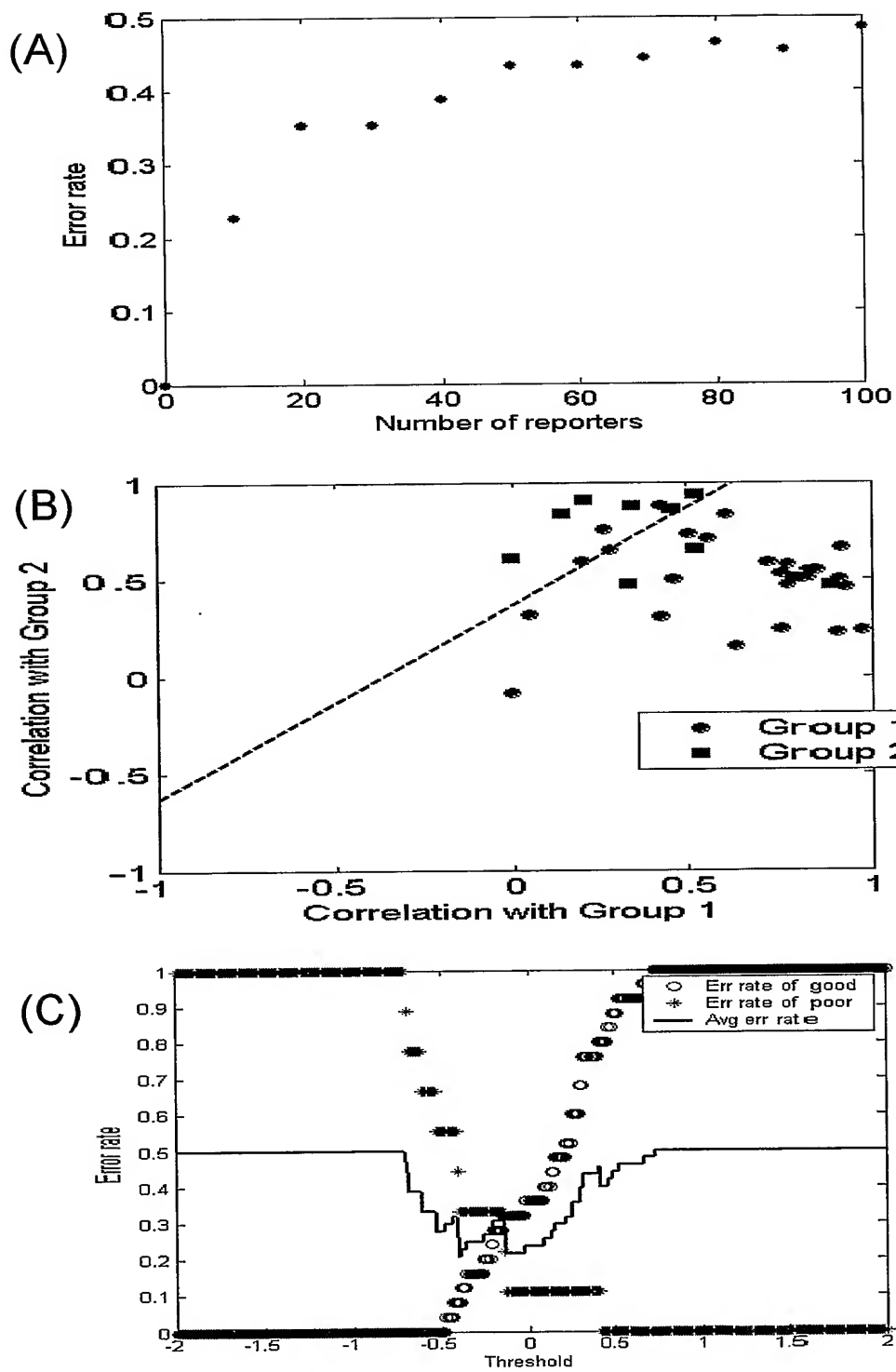
FIGS. 4A-C



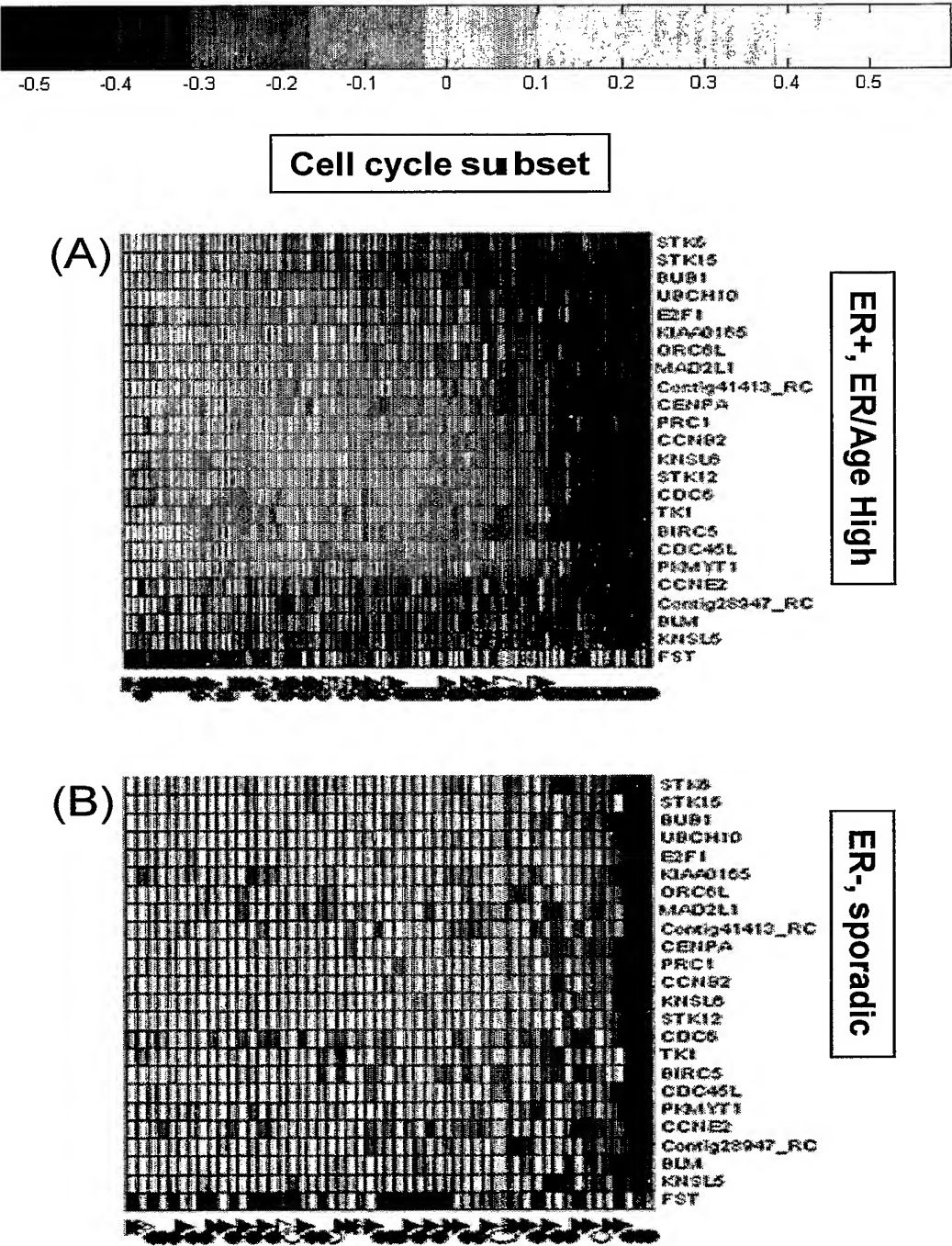
FIGS. 5A-C



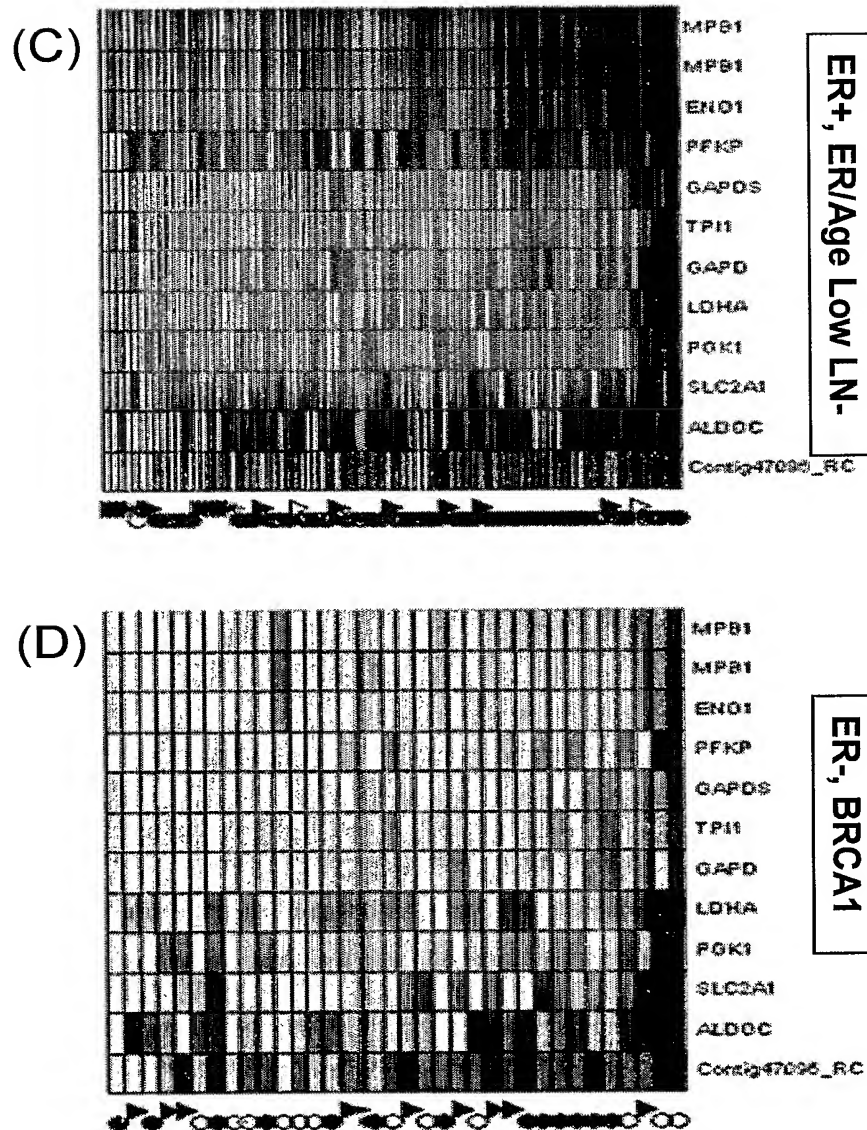
FIGS. 6A-C



FIGS. 7A-C



FIGS. 8A and 8B

Glycolysis subset

FIGS. 8C and 8D

Sequence Listing

<110> Rosetta Inpharmatics, LLC

<120> Classification of Breast Cancer Patients Using a Combination
of Clinical Criteria and Informative Genesets

<130> 9301-251-228

<140>

<141>

<150> 60/650,401

<151> 2005-02-04

<150> 60/604,076

<151> 2004-08-24

<150> 60/550,810

<151> 2004-03-05

<160> 366

<210> 1

<211> 4946

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 1

cagcctcagc	ccccagatga	agatggggat	cacagtgaca	aagaagatga	acagcctcaa	60
gtggtggttt	taaaaaaggg	agacctgtca	gttgaagaag	tcatgaaaat	taaagcagaa	120
ataaaggctg	ccaaagcaga	tgaagaacca	actccagccg	atggaagaat	catatatcga	180
aaaccagtca	agcatccctc	agatgaaaaa	tattcagggt	taacagcaag	ctcaaaaaag	240
aagaagccaa	atgaagatga	agtaaatacag	gactcgggtca	aaaagaactc	acaaaaacaa	300
attaaaaata	gtagcctcct	ttcttttgac	aacgaagatg	aaaatgagta	agtgtaaata	360
ttttgaattt	agtctacttt	gaaagtataat	ggagtgttca	ttaaaatcac	atTTTTTcct	420
attataaaga	tactacaagt	tctttataga	aagtttagga	aatagagaaa	aaaatttaata	480
aaactacatc	tattcatcaa	taccctctcg	acttaaaatg	ccaactctat	agaaattagc	540
tagtattaac	atTTTgttat	ttcccttgTg	tggttgTata	tatatgtaaa	ttatatTTTT	600
aagcaaaata	cattTTTTgt	gtgtaaacaa	aattttataa	atacaactgt	attgcaaatg	660
ttctttgtcc	tgtctctcac	ttgacattgc	attatgagta	ttcttccagg	tcagttaaatt	720
tcaaaaacct	gacattaata	gctacagata	atttcataaa	catctcattg	tatctTTTTc	780
attagcaata	gctccacttt	gggtggggga	gatgataatg	tgccttgTta	aaaatacctc	840
cccaactcct	gctaagggtg	gccatgagac	tcagctctgg	caagttaaga	aatacagggtg	900
gaattctgct	tgataaagct	gctgggtttt	ttgttacaaa	aggacagact	tggcaaacat	960
gagcctttgc	tcttatcttt	tcatocta ct	tggagtgcag	agataaaaacc	tgagtaccag	1020
agccactttt	aggcataagg	aaggcagcca	tgtgctttgg	gtcatgttag	taaaaagact	1080
cagagctttg	ctccttgctg	acatgcctgg	aggagctgct	acaccagctt	ggattgctga	1140
cctctgactt	cttggttagtg	agaagaataa	acactgtgct	taattaggcc	ttggtcagggt	1200
ttcttttata	tgcagccaaa	tgcagtcc ta	agtaatacaa	taaataactg	gtcaaaactgt	1260
tactggtgga	gggtgtccag	gttcttggca	ttttggacaa	ataattgaac	aaaacgcaca	1320
aagcaatgaa	tatcctctag	aggtttgc ca	ttggttactt	ggcgtacacc	ctgtgtaaat	1380
gaagtagtgg	cccgtgacct	gtctgattgg	tgcagaaagt	gaccaatcag	aggctgaagt	1440
gaagttacaa	agttatactc	ctgtgtaaat	gaggacttgg	cctatgacca	gtctgattgg	1500
ttgcaggagg	ggaccaatca	gaggcaactt	cattttttcat	ctgcaatgca	gaaaaggcaa	1560
ggggattgca	aagggtagtag	cctctgatcc	ttttgttact	taggtatgga	gagggtgggt	1620
tttccttttg	attcagttct	aggaagtcaa	tgtgaatcag	ccttaggttc	cctgtctcca	1680

gaccctat	tcctgcct	ttttcccc	gagagac	atcctcgt	atctttat	1 740
gaggctga	gactgagg	ctttcttc	taactgct	atgctaact	gggacacag	1 800
ccctacct	tggagatc	gtaactct	ccctgctt	tctaggggag	acagggtagc	1 860
ttcttgat	ccgtggtg	cttctcct	aactggct	aaatcttg	acatgatcat	1 920
ctaacttg	ggtctctag	caaaaggaa	tggatttgg	taaaagatt	aacagatag	1 980
gtccaaaa	caaggcaa	ataatcat	ataatggg	ggccaagg	gggagccat	2 040
aaaccca	tagtgcc	taggtgcc	agctgttg	atattttag	ggcccagtc	2 100
gctagttt	agtggtgt	ccttacta	cctgattg	tgacatcaa	acagcattct	2 160
tcttctag	aaatacata	gccacctg	tcagcagtt	ggagatctag	ttcccctcga	2 220
ttttgcaa	cgaccact	caaggagc	atccgaatt	gtaagggtg	aatactttga	2 280
gcaatggt	ccaggcttt	cataaaat	ttggacaag	gttggtata	ggatagggaa	2 340
gttgcaat	cgctaact	cattccta	tctgctgt	ttcctagcc	ttgtgtctg	2 400
tgggtgcag	taaaggata	atgaggga	gggtgtgg	agctatatta	atttagggac	2 460
atacaatat	tctgtctcca	gtctacca	ttccaccaa	acaaatcaca	gcagaaccga	2 520
cctaactt	aaataaact	cagtccca	tactgggc	gattaccac	acaaagtgc	2 580
acaagaat	ttgtccat	agactctc	agattggc	tgctagaaca	tttcacaagg	2 640
ccatttcag	caaagtc	agaaagta	cggtttca	tgtgccctat	tacaaaagaa	2 700
aacgtgtt	ttaacttt	acagacaa	gccatgaat	aagaatatt	ataaatagtt	2 760
tacaaatt	ggagaaat	gaatactc	tacacttaa	gtgtatttc	aggctataaa	2 820
tagctcaaa	taaaaag	attcagac	tgaaaaaa	aaaagaag	gcaatatttc	2 880
aaacaacaa	agccataca	attatttc	tcttccatt	gttcatttc	gtccatgtaa	2 940
tcaactcct	ctctacttc	tattcatc	tatgaacac	tcagccttc	aattagtgc	3 000
ttggaagtt	tctgtctaat	ccaatggc	actctccaa	gttaccaga	acctgcatt	3 060
aagagttct	ttcatgaac	ccaaagaag	aagccttgg	ctgtagctga	ttataagtc	3 120
ctttttttt	ttgagaagg	tcaaagcaa	acatcaatta	tggatgacaa	aagtcttaag	3 180
acagccata	agacacagt	gacaaatg	gctatttct	tggcttacaa	caatttaaca	3 240
taatcattac	aacatatatt	aagacata	agaattttg	aactctcata	caatcctgga	3 300
acacatatta	acaacaaat	tctatcag	taacccaa	gaagctaaac	accacctcac	3 360
acttgacaat	gtttcctgt	taattcaa	attacaaat	agcctaatat	aagcctaata	3 420
tgtcactct	gaacttcag	aagcctaata	tccaaaaag	tagtttaagg	tcaaaagttt	3 480
ttgaattaac	ttttttccat	tagtatgg	atatctttt	tactaatttg	taagttatgt	3 540
aatttatcaa	tttttttttg	ttgttctgt	tcccaacct	tatgtcagat	aaagaatcac	3 600
ccaggccaga	cacagtggc	catgcttgt	gtcccagc	tttggaagc	caagggtgga	3 660
gaattgcttg	aagccaggaa	tctgagcca	gcctgggcga	caaagcaata	cccctatctc	3 720
tacaaaaaat	aaaaaatag	caggtgtgg	gacacacac	tgtggtccca	gctgctcggg	3 780
aggctgagcg	ggaggatgg	ttggggccag	gggttcaac	ctgcagtga	ctgtgattgc	3 840
gccactgcac	tcagcctgg	gcaacagag	aagaactgt	tcaaaaaaa	taaaaaatag	3 900
aaataaattt	taaaaaaaga	attaccata	ttctctttg	ttttgtttat	tcacattaac	3 960
ctttattcta	tctggaatt	atttgagtat	acttttttct	caaataatca	attgtcctag	4 020
aacctatgt	ttctcattta	tttgaaagg	catctagtga	gagatttctc	caaagtgttg	4 080
ggtagggga	ggagggaag	cacttttaag	tctgagcct	tagagggtgat	tcctcaagac	4 140
cctgcttaat	cctaacaatt	ttctctatta	gtaaaagtca	gcccaaactg	ggggcctgtt	4 200
aagatcctta	ccagccacat	ccatctgaaa	ttatgaattt	caaagtatct	tacaaatttg	4 260
gtgccacatt	atctttttta	agtttgtttt	gttttgtttt	tttgagacag	agtctcgctc	4 320
tgtcaccccg	gctggagtgc	agtggcg	tctcagctca	ctgcaagctc	cgcctcctgg	4 380
gttcacacca	ttctcttgcc	tcggcctccc	aagtatctgg	gactgcagtc	gcccgcacc	4 440
acgcccggct	aatttttttg	tatttttaaat	agagacggg	tttcactttg	ttagccagga	4 500
tgggtctcaat	ctcctgacct	catgatacac	ctgcctcggc	ctcccaaagt	gctgggatta	4 560
caggcaggag	ccaccgcgc	tgggcctttt	tttaagtttt	aagtacctat	aaagaacact	4 620
gaaagggtgat	gtgtgtggat	gagctaggaa	gacctgaaat	aggctctctc	taaattaatc	4 680
aaattaatcc	tgaagccatt	ctgcaatact	gtcttttaatg	tatactcact	tgttatagaa	4 740
gccagggttt	tttcccctaa	ttgtatcat	tgtctatatgt	gttattgtac	caaactacac	4 800
tgttttaatt	gctgtaaatt	ttaatatgtc	ttagtatctg	ggtgtgggaa	tcttgaaagc	4 860
atggagtttg	tgttattcac	cactgtattc	tcaaatatca	gaagagtatc	tggcctacta	4 920
agtgcacaat	aaacatagtt	aaaatg	4946			

<210> 2

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 2

taatcctgaa gccattctgc aatatgtct ttaatgtata ctcacttggt atagaagcca 60

<210> 3

<211> 1007

<212> DNA

<213> Homo sapiens

<300>

<308> AF005487

<400> 3

gaatacagaa	tgtgggcaaa	ctcgcttctg	tgccggccgc	cagaaggttt	gctgagggca	60
atcactccct	ggtgccgggc	tccttgaggt	tatgcactgg	gacatctaga	gcctattggt	120
tgaggaatgc	agtcttgcaa	gcctgctctg	gatcaagcca	cagactgaaa	cacccccgaa	180
gagcaagcac	gtttcttgga	gcaggcctaag	tgtgagtgtc	atatcttcaa	tgggatgaag	240
cgggtgcagt	acctgaacag	atacatccat	aaacgggagg	agaacctgcg	cttcgacagc	300
aacgtggagg	agttccaggc	agttacggaa	ctggggcggc	ctgtcgcaga	gaactggaac	360
agccagaagg	gcatcccggg	ggagraagcgg	gacaagatgg	acgactactg	cagatacaat	420
tacgggggttt	tttgagagct	tcacagtgca	gccgcgagtc	catcctaagg	tgactgtgta	480
tcctgcaaag	acccagcccc	tgcatcaccg	caacccctcg	gtcggctctg	tgagtgggtt	540
ctatccaggc	agcattaaag	tcagggtggtt	ccagaatggt	caggaagaga	aggctgcggt	600
ggtctccata	ggcctgatcc	agaatggaga	ttggaccttc	cagaccctgg	tgatgctgga	660
aacagttcct	cggagtggag	aggtttacac	ctgccaaagt	gagcatccaa	gcgtgacgag	720
ccctctcaca	gtggaatgga	gtacacggac	tgaatctgca	cagagcaaga	tgctgagtgg	780
agtcgggggc	tttgtgctgg	gcctgctctt	ccttggggaca	gggctgttca	tctacttcag	840
gaatcagaaa	ggacactctg	gacttcagcc	aacaggactc	ctgcgctgga	ctcctgagct	900
gaagtgcaca	tgaccacatt	caagggaagaa	ccttctgcca	cagctttgca	ggatgaaaag	960
ctttcccaact	tggtctcttat	tcttccacaa	gagctctctc	aggacca	1007	

<210> 4

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF005487

<400> 4

tttgcaggat gaaaagcttt cccacttggc tcttattctt ccacaagagc tctctcagga 60

<210> 5

<211> 3200

<212> DNA

<213> Homo sapiens

<300>

<308> AF026941

<400> 5

caggaagggc	catgaagatt	aataaagatt	tggactcagg	gcaaataatt	acttagtagc	60
aataactcaa	agaattactg	ttgaataaat	aagccaatta	agcagccaat	cacgtactat	120
gcggatgcac	acaaatgaaa	ccctcacttc	aacctgaaga	cattcgcaca	tgagttacgt	180
agaggggacct	gcaggaagcg	gtagagaaaa	cataaggctt	atgcgtttta	tttccacacc	240
aatttcagga	tctttgtcac	tgacagcagc	actaagactt	gttaacttta	tatagttaag	300
aagaacaagg	ctgagcgcga	tgactcacgc	ctgtaagcct	agaacttttg	gaggccaaag	360
caggcagact	gcttgagccc	aggagttcca	gaccagcctg	ggcaacatgg	caacacccca	420

tctctacaaa	aaaatacaag	aatcagctgg	gcgtgggtgat	gtgttcctgt	aatctcagct	480
actcgggagg	cagaggcagg	aggatttgctt	gaaccccgga	ggcagagggtt	gtagttagcc	540
gagatctcgc	cactgcactc	cagtctggac	gacagagtga	gactcagctct	caaataaata	600
aataaataca	taaatataag	gaaaaaaata	aagctgcttt	ctcctcttcc	tcctctttgg	660
tctcatctgg	ctctgctcca	ggcatctgcc	acaatgtggg	tgcttacacc	tgctgctttt	720
gctgggaagt	tcttgagtgt	gttcaggcaa	cctctgagct	ctctgtggag	gagcctggtc	780
ccgctgttct	gctggctgag	ggcaaccttc	tggtgctag	ctaccaagag	gagaaagcag	840
cagctgggtc	tgagagggcc	agatgagacc	aaagaggagg	aagaggaccc	tcctctgccc	900
accaccccaa	ccagcgtcaa	ctatcacttc	actcgccagt	gcaactacaa	atgcggccttc	960
tgtttccaca	cagccaaaac	atcctttgtg	ctgccccttg	aggaagcaaa	gagaggattg	1020
cttttgcctta	aggaagctgg	tatggagaag	atcaactttt	caggtggaga	gccattttctt	1080
caagaccggg	gagaatacct	gggcaagttg	gtgaggttct	gcaaagtaga	gttgcggtctg	1140
cccagcgtga	gcatcgtgag	caa.tgggaagc	ctgatccggg	agaggtgggt	ccagaattat	1200
ggtgagtatt	tggaacttct	cgc.tatctcc	tgtgacagct	ttgacgagga	agtcaatgtc	1260
cttattggcc	gtggccaagg	aaagaagaac	catgtggaaa	accttcacaa	gctgaggagg	1320
tggtgtaggg	attatagaat	cccttttcaag	ataaattctg	tcattaatcg	tttcaacgtg	1380
gaagaggaca	tgacggaaca	gatcaaagca	ctaaaccctg	tccgctggaa	agtgttccag	1440
tgctctttaa	ttgaagggtga	gaattgtgga	gaagatgctc	taagagaagc	agaaagattt	1500
gttattgggtg	atgaagaatt	tgaaagattc	ttggagcgcc	acaaagaagt	gtcctgcttg	1560
gtgcctgaat	ctaaccagaa	gatgaaagac	tcctacctta	ttctggatga	atatatgcgc	1620
tttctgaact	gtagaaaggg	acggaaggac	ccttccaagt	ccatcctgga	tgttgggtga	1680
gaagaagcta	taaaattcag	tggaatttgat	gaaaagatgt	ttctgaagcg	aggaggaaaa	1740
tacatatgga	gtaaggctga	tctgaagctg	gattggtaga	gcggaaagtg	gaacgagact	1800
tcaacacacc	agtgggaaaa	ctcctagagt	aactgccatt	gtctgcaata	ctatcccgtt	1860
ggtatttccc	agtggctgaa	aacctgattt	tctgctgcac	gtggcatctg	attacctgtg	1920
gtcactgaac	acacgaataa	cttggatagc	aaatcctgag	acaatggaaa	accattaact	1980
ttacttcatt	ggcttataac	cttgttgtta	ttgaaacagc	acttctgttt	ttgagtttgt	2040
tttagctaaa	agaaggaat	acacacagga	ataatgacct	caaaaatgct	tagataaggc	2100
ccctatacac	aggacctgac	atttagctca	atgatgcgtt	tgtaagaaat	aagctctagt	2160
gatatctgtg	ggggcaatat	ttaatttggg	tttgattttt	taaaacaatg	tttactgcga	2220
tttctatatt	tccattttga	aaactatttct	tgttccagggt	ttgttcattt	gacagagtca	2280
gtattttttg	ccaaatatcc	agataaccag	ttttcacatc	tgagacatta	caaagtatct	2340
gcctcaatta	tttctgctgg	ttataatgct	tttttttttt	tttgctttta	tgccattgca	2400
gtcttgtact	ttttactgtg	atgtacagaa	atagtcaaca	gatgtttcca	agaacatatg	2460
atatgataat	cttaccgaat	ttcaagaagt	ctctagaaag	agataacaca	tggaaagacg	2520
gcgtggtgca	gcccagccca	cggtgcctgt	tccatgaatg	ctggctacct	atgtgtgtgg	2580
tacctgttgt	gtcccttttct	cttcaaagat	ccctgagcaa	aacaaagata	cgctttccat	2640
ttgatgatgg	agttgacatg	gaggcagtg	ttgcattgct	ttgttcgcct	atcatctggc	2700
cacatgaggc	tgtcaagcaa	aa.gaatagga	gtgtagttga	gtagctgggt	ggccctacat	2760
ttctgagaag	tgacgttaca	ctgggttggc	ataagatatc	ctaaaatcac	gctggaacct	2820
tgggcaagga	agaatgtgag	caagagtaga	gagagtgcct	ggatttcatg	tcagtgaagc	2880
catgtcacca	tatcatattt	ttgaatgaac	ctcgagtcag	ttgaaatagg	gtaccatcta	2940
ggtcagttta	agaagagtca	gctcagagaa	agcaagcata	agggaaaatg	tcacgtaaac	3000
tagatcaggg	aacaaaatcc	tctccttgtg	gaaatatccc	atgcagtttg	ttgatacaac	3060
ttagtatctt	attgcctaaa	aaaaaaatttc	ttatcattgt	ttcaaaaaag	caaaatcatg	3120
gaaaattttt	gttgtccagg	caataaaaag	gtcattttta	tttaaaaaaa	aaaaaaaaaa	3180
aaaaaaaaaa	aaaaaggcca	3 200				

<210> 6

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF026941

<400> 6

atttttgaat gaactctgag tcagttgaaa tagggtagca tctaggtagc ttttaagaaga 60

<210> 7

<211> 1799

<212> DNA
<213> Homo sapiens

<300>
<308> AF035284

<400> 7
gcttgaaccg gggaggtgga ggttgcagtg agctgagatc acgccattgt actccagcct 60
gggcgacaga gcaagactcc atttcaaaaa aaaaaaaaaa aaaaaaaatc cactcatata 120
aaaggtgagc tcagctcact ggtccatttc tcagtggctt ctccatcctc atttgcaaac 180
ctcagagggga taaggcagtt gaacctgatg agcaagaatt ataacagcaa ggaaacatta 240
atgcttagaa ttctgagatc cagcacaact cagtctgtgg gagctcagct cgctgccag 300
ggataggtat gacctatgtc tgcccttaggc tgctgggaga tgccattctc cagtttcaga 360
agcaggcagg gcaaagggtca agactgtggt attgggggtct tttggctctg aaggatcctg 420
gaaccactga ttttggttta ttccctccag ggtctaaaga gaacaagagg tgctagctct 480
taccaaaaca gatggtagag agagtgtctg gctattttaa aagctctttc atcttttaat 540
tcacctcttc ttttcacctc tttaaccact cctcaggaac agaacttctc taggactggg 600
ggtcttttag ctccataagc aagtgcagcag atgggacaag ttagtctttt ctccctagaa 660
acaaagggga tgcccagtggt ttcccttttg ctcccaacc taaaatttca agtttaataa 720
aatagcaatt agcagaagtg accaaattgg gagataatta tcagtcatga ggaaagacac 780
agatttcggg cataaagaat gtaagggtta taagtagaaa ctttctataa cctaaatgat 840
gttatagaat tatttttgag caggagcaga aagattaaat atgatcactt catacttcta 900
aatcagaaat aggaagatta aaaccacaga acagtttgtg atttctattg ctggtagcta 960
ggtatcttac tctgtccact ctgtttcaag tatctaactc ttctggaaac caaataggct 1020
ttagaagaga ttatcctata ttccctatcag tataatacta aaatgtaact ttttaatcat 1080
ctggttttta aaagataaac agtttagccc atctctccag agagcaaaca taggaatatg 1140
actcaggagc ctccatagggc ttatcatcag cctcacacc cgcttcccc tccaaccac 1200
agcctttgct tccaggtggc aggattacta ctttgctctc tcagcagcat ctactctagg 1260
catattgatc atttttagaca ctgggagaag agaacctcaa actaggagga aaagacagag 1320
cctccactta gttttgggag gggatggcag acagtcgaagg agatgagcgt cctaaggcat 1380
gttgggatag ggtcagatgc accacccatg gagaggtttg tcaacacaaa gacatggaag 1440
gtttagagggt tgtcaacaaa aagacatgga aggttaggtt tgtcaacaca aagacatgga 1500
agatttagagg tttgtcaaca caaagatata ggaagaatgg gctgcagaag atttagatgt 1560
tttccatttg ggcacatttt acctagctgg agaactaggt ttaaaacagc ctgggtagga 1620
aaattagaag caagctggat gcagtggctc atgcctgtaa tcccaacact tttgggaggt 1680
ccaggcagga ggatcacttg ggcccaggag gtcaagcctg cagcgagctg agatcacacc 1740
actgcactcc agcctggggt gataagaaca gaccctgtct caaaaaaaaaa aaaaaaaaaa 1799

<210> 8
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> AF035284

<400> 8
caaaaagaca tggaagggtta ggtttgtcaa caciaagaca tggaagatta gaggtttgtc 60

<210> 9
<211> 1380
<212> DNA
<213> Homo sapiens

<300>
<308> AF052162

<400> 9
gtcaaaggat atttatattat aggccttttt ttttttaata tagaatctga ggctgtttgg 60
gctttgactt aaattttccat caggcctctc tccagcaggt aatccctctc cttccgctgg 120
gtcccctggg gaggtgtgaa ctcaagggcc tagcccaaaa acactttttc tgctttttct 180

aatcccttttc	cagtccectc	tttttttata	aacgtttggca	gtttgatgtt	tctgtttcgg	240
cataacgtaa	tccatttcac	tgtagcctaa	actccagtc	gaggttgat	attgttcaaa	300
tgagcagggc	ccgagctgga	agcgcaaggc	agccgccgcc	gtgccgctcc	tcccttgccc	360
tcaggccagg	tccctgctgg	aagcggctgc	atcttcctgt	cagccctggt	ttccatgggtg	420
actggcgtca	cgcagccacc	cgagtatggc	tgaccttcct	gcagagagag	gagccgcagt	480
cttttgcttg	tggaaggaga	cgctgggctg	tgcggtgcgg	aggggtgatga	ggatgtctgg	540
tgacagccgt	gcggacacca	ctcctctctg	cagcactgcc	tcccagcgcc	agggtcgcgg	600
gcacatccca	ctgagagcgg	gggtcctgcc	ccatcttaga	gtcaaaggca	gaggggcttc	660
caggccctgg	atgggggtatt	ttgggtgtcac	ctgaagtccc	tctgacatca	ccttgtttca	720
tcatttttta	tgacagaatt	agaaacccat	ccttcaagca	caataatcat	cacagacttg	780
agtttgcttc	ctaaagcaaa	ggctccgggt	ttgtttggaa	aatttttttg	atttctgaaa	840
tgaattgatt	tttatatttg	gggcatctct	atagaaagtg	accaccaagg	ccagtaagta	900
cgggaaaaaa	tgtttactaa	cttcctcaga	gattcgtgat	acgcgtttct	ccactgacag	960
acatttaaaa	acaaccttca	gctccgtttc	aatcaatcac	ctcgacttgt	tttttagcat	1020
ggacactgcc	agcaggacag	acagggatgg	agtaaaccga	agtcaatttc	agggctcttg	1080
gcgtgttgga	cacagaagaa	atcctagtgc	agcctttggg	agctaacagt	caactgatttt	1140
ataattggag	aatgcgtaaa	gattcatttt	tcaaggagaa	gagcctgcaa	atggccaatg	1200
aaggaggtaa	ataaactaag	atattccgag	ggaagggacc	caggccacct	cccttcogca	1260
ggtctgcaga	tgaagggttt	tttgaatgaa	atgccactgt	gcattttcag	aaaaaaaaat	1320
ctctgataaa	cagactttga	atggaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1380

<210> 10

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF052162

<400> 10

cagtaagtac gggaaaaaat gtttactaac ttccctcagag attcgtgata cgcgtttctc 60

<210> 11

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 11

ggggaaaaa	gctaggaaa	agctgcaaag	cagtgtgggc	tttttccctt	tttttgctcc	60
ttttcattac	ccctcctccg	ttttcacctt	tctccggact	tcgcgtagaa	cctgcgaatt	120
tcgaagagga	ggtggcaaa	tgaggagaaa	gaggtgttag	ggtttggggt	ttttttgttt	180
ttgtttttgt	tttttaattt	cttgatttca	acattttctc	ccaccctctc	ggctgcagcc	240
aacgcctctt	acctgttctg	cggcgccggc	caccgctggc	agctgagggt	tagaaaagcgg	300
ggtgtatttt	agattttaag	caaaaaat	aaagataaat	ccatttttct	ctcccacccc	360
caacgccatc	tccactgcat	ccgatctcat	tatttcgggtg	gttgcctggg	ggtgaacaat	420
tttggtggctt	tttttccctt	ataattctga	cccgtcagg	cttgagggtt	tctccggcct	480
ccgctcactg	cgtgcacctg	gcgctgccct	gcttccccca	acctgttgca	aggctttaat	540
tcttgcaact	gggacctgct	cgcaggcacc	ccagccctcc	acctctctct	acatttttgc	600
aagtgtctgg	gggagggca	ctgctctacc	tgccagaaat	tttaaaacaa	aaacaaaaac	660
aaaaaaatct	ccgggggccc	tcttgggccc	tttatccctg	cactctcgct	ctcctgcccc	720
accccgaggt	aaagggggcg	actaagagaa	gatgggtgtg	ctcaccgcgg	tcctcctgct	780
gctggccgcc	tatgcggggc	cggcccagag	cctgggctcc	ttcgtgcaact	gcgagccctg	840
cgacgagaaa	gccctctcca	tgtgcccccc	cagccccctg	ggctgcgagc	tggtcaagga	900
gccgggctgc	ggctgctgca	tgacctgcgc	cctggccgag	gggcagtcgt	gcggcgtcta	960
caccgagcgc	tgcgcccagg	ggctgcgctg	cctcccccg	caggacgagg	agaagccgct	1020
gcacgccctg	ctgcacggcc	gcgggggttt	cctcaacgaa	aagagctacc	gcgagcaagt	1080
caagatcgag	agagactccc	gtgagcacga	ggagcccacc	acctctgaga	tggccgagga	1140

```

gacctactcc cccaagatct tccggcccaa acacacccgc atctccgagc tgaaggctga 1200
agcagtgaag aaggaccgca gaaagaagct gaccagtc aagtttgtcg ggggagccga 1260
gaacactgcc ccccccgga tca tctctgc acctgagatg agacaggagt ctgagcaggg 1320
cccctgccgc agacacatgg aggccttcct gcaggagctc aaagccagcc cagcatggt 1380
gccccgtgct gtgtacctgc ccaattgtga ccgcaaagga ttctacaaga gaaagcagtg 1440
caaaccttcc cgtggccgca agcgtggcat ctgctggtgc gtggacaagt acgggatgaa 1500
gctgccaggc atggagtacg ttgacgggga ctttcagtgc cacaccttcg acagcagcaa 1560
cgttgagtga tgcgtcccc ccaaaccttt cctcaccgcc ctcccacccc cagccccgac 1620
tccagccagc gcctccctcc accccaggac gccactcatt tcattctatt taagggaaaa 1680
atatatatct atctatttga gga aaaaaaaa aaaaaaaaaa aa 1722

```

<210> 12

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 12

```

tccaccccag gacgccactc atttcatctc atttaaggga aaaatatata tctatctatt 60

```

<210> 13

<211> 1411

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 13

```

aaacaaagag atgccacccc tgtgtgatgg ctttgggtacc cgaacactga tggttcagac 60
attttcccgt tgcattctgt gttccaagga tgaagtggac ttggatgagt tattagctgc 120
tagattggta acgtttctga tggacaatta ccaggaaatt ctgaaagtcc ctttggcctt 180
gcagacctct atagaggagc gtgtggctca tctacgaaga gtccagataa aatacccagg 240
agctgatatg gatatactt tatctgctcc atcattttgc cgtcaaatta gtccagagga 300
atttgaatat caaagatcat atggctctca ggaacctctg gcagccttgt tggagggaagt 360
cataacagat gccaaactct ccaacaaaga gaaaaagaag aaactgaagc agtttcagaa 420
atcctatcct gaagtctatc aagaacgatt tctacacca gaaagtgcag cacttctggt 480
tcctgaaaaa ccaaaccga aaccacagct gctaattgtg gcactaaaga agcctttcca 540
accatttcaa agaactagaa gttttcgaat gtaataatac ttccacagca acaggtgcta 600
gagaccactg ttgttggttt gagtgaatgg tggtaggag aaagactttg gtggtggaag 660
aaagaaaagc ataaaacaaa gactactgaa atatagataa agattgcctt agtttttaaa 720
aatgtttggc cattagtatt ttataaaaac tcaatgctag ttttaagtgt ataaattggt 780
taaaatttat gagtcaaata tatagtata atgttaacat gtttgtaatt gctacagaat 840
ttaagggtat tttatctct gtgctttctt tttcatgggt tttattaaat aattgtgtat 900
atacatccta gctactgata tctttattat agccttaaga ctttaattta agtcttaaaa 960
atagcgtgta tacttgaata agaaagacac tgggtactgt tactgtgatg ctattgactt 1020
agtagccaat tatcatttct cctgtataaa ttccagtttt tattgtgcga cataaatttt 1080
ttaatgtctt atattgtgat agctatgtct tttattgcag atttattgga tgttatgaca 1140
gattttacta aagctagtgt ttttataaca tatatattag ttgatgttta cctataagtg 1200
gagtagattt tcactcgctt gcaatgggat aatttcagtc ttagctaaaa atggaaagtt 1260

```

```

gaactggata aattcttttg gtacccttag acctctgatt ctaagtcaaa tgcaaattggg 1320
ttaaataaaa tgagactact tcttttataa atataatttc atccttttga aagtaagtga 1380
aatgtaaata aacttatttt ttttaaaaat g 1411

```

<210> 14

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 14

acccttagac ctctgattct aagtcaaagtg caaatggggtt aaataaaatg agactacttc 60

<210> 15

<211> 2352

<212> DNA

<213> Homo sapiens

<300>

<308> AL049367

<400> 15

```

ggcaaacccc ttttaaaatc taatgtctgg gctttgagta ttagctcatt taggggtggac 60
aaatgcatta ctgttttcaa actgctcaca tttattcagt atttctccaa gttgctatct 120
actcagcctt atgaatgcc ccctcgcttttc taaggccatg tgaaaatcac ggcactgccc 180
ttagccttgt gtcactctgct ttttcgttct gogatatgcc cagttccaa atcaattata 240
ggtacctgtt taggagagag gaagatttta cctctcaaag ggtgagattt gaaatttaca 300
ctaaaaagac aactttacat ttaatgcttc acttaatgag acattctttt ttttataaagt 360
ctatttttct actcagtttc agaacactaa tctgattttc actctgattt ttaacgtttc 420
tttaaatatt tataatgtag ctcttttcaa aatattttca tgaaaaatta cttttaatat 480
accattatgt gcatgttatt ggtagcaggc atagtttatt atttagtact gaaacatgct 540
cttttaccta acagtaaaaca agtatgtttt gatataatc tgtaaatatg cttatagtgg 600
taagaaatgg acttgagggtc ccaggagatt tcattttatt caccctgggtc agatacaata 660
aaggctatga gtataaatac ataacttcct aaccagggtgt agggcatgtt catgaatata 720
aaatcttttg atgttggaac caagagagga aaagttgtag ctaaatgttg atttacttat 780
aactagacgt ctatgtgaga aaatatatgt atacatatat atgatatgca gaagtcactt 840
tttttatcag gctttattct ccttacaag ccacagttta actgtctgca acagttgggt 900
tatgttaatg atagacaaat acccagtggt tgttactttt tccaactacc actgtaatga 960
taatctttct cacgtatata catgcaactt cttggcttca tttccatgaa gctgttcaa 1020
tatattcagt atactttgtc cttaatgctg cttctgttaa cagtgatctc tttctttt 1080
tcattcttat atcttcatca gttcatcata aatctgtcca gttgaggcct caggacacag 1140
gcatgatttc atgactcga agtattttac agaaacattt tttaaataag ggaaatattt 1200
tatataccag atggttcaca agtgatggct catagctagt tttttttttt tcttctaaaa 1260
aatgtcaggt ttttaaaatc atttacctta ttaaaatgaa aagtgccata cttaaccttt 1320
aaaggaaaga cctgacttgc tttttctcta tttagactgt ttttgtactt tactaaactt 1380
taaactatca ggaaaaaac caaaacttta taccaatgat ttagtaattt tgaggcatag 1440
ggtagcttac gtatggagg atgtgcaaaa tattctcttc aaatgccacc ttctcaattt 1500
ataactaaaa tagtgttatc tgactaatc ctctgaattt tgatgtaaga tctataagg 1560
ccccaaaat gatcgtagta catgccagtc atttctcagt gaaataaata caatacaga 1620
gtacattatg ggttttatg ctttctttta tggtagacct gttaatgggg aaaaaatata 1680
tcaaatcaaa tagaatctta tatctgtatg ttaaaataga gcacttacct gaagtcagt 1740
gcctggatca tagccctgga tcatttccca gtctgtcctg tgctgtgtga ccttggacaa 1800
ggcgcttcat ctctctgggc ctctatttct ccatttgtaa aacaagtggc tgcagtatag 1860
gatggctgag agcccttct gttcccagat gccttgggtc aaagacccca cccctctgct 1920
ggtcctgcca acgtgttgg gctataagct gcttcagata taaaattgggt ttatctataa 1980
tgtttgttca ttaataagct tctaaaaggc ctttttgtta tacagtgtt tttttctagt 2040
tttatggact tgattactgt aataatgtct tgttttttagc catgtaacta caaacagata 2100
ttctcttgat gcttagttaa atttgcattt gatatatcat tgatgagatt ttgttgttat 2160
gtaattattct ttggctacgc atctgtccag catcttatta accataatac tgtgatcatt 2220
atgtggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgtcac 2280
agatttgaaa gaagtttcat taaaagcacc attgctttct gtaaaaaaaa aaaaaaataa 2340
aaaaaaaaaa aa 2352

```

<210> 16

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL049367

<400> 16

atttggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 60

<210> 17

<211> 1130

<212> DNA

<213> Homo sapiens

<300>

<308> AL080235

<400> 17

```

ggtcgcgcga ccggccgcct ccggcccgcc gccgccccca gccgcgcgcg cgcacccgcc 60
ggggcgccca ccgcgctgcc agcctacccc gccggccgagc cgcccgggcc gctgtggctg 120
cagggcgagc cgctgcattt ctgctgccta gacttcagcc tggaggagct gcagggcgag 180
ccgggctggc ggctgaaccg taagccattt gagtccacgc tgggtggcctg cttcatgac 240
ctggtcatcg tgggtgtggag cgtggccgcc ctcatctggc cggtgcccat catcgccggc 300
ttcctgcccc acggcatgga acagcgccgg accaccgcca gcaccacgc agccaccccc 360
gccgcagtgc ccgcagggac caccgcagcc gccgcgcgcg ccgcgcgtgc cgcgcgcgc 420
gccgcgcgtc cttcgggggt gccgaccaag tgaccgcgtc cgctcctccc tgtgtccgtc 480
ctgtgtccgc gcgcgcgggt gcccttcccc cgggggactc ggccgggtgt cttcgtgctg 540
tagttatcgt tagttcctct tcccagatg gggccgcca gagacccag cgcctttgaa 600
aagcaagggt tgtgctgcgc ttccagttcc gaaaagcaga tgtttaagcc cttggactga 660
gggtgggata gcagctccga agacggagag gagggaaatg gggccctttc ccctctattg 720
catccccctg cccgactcct tccccgcacc cagtgccct agattcatgg cagaaaatga 780
ccaaatcctg tgtatttgtt ttatatattt aataactgtt ttaaatgaaa gtttttagtaa 840
aaaaaataca aaacaaaaag attaaattgc tattgctgta gtaagagaag ctctttgtat 900
ctgaacatag ttgtatttga aatttgtggg tttttaattt atttaaaatt ggggggaggg 960
catgggaagg atttaacacc gatataattg taccgctgaa aatgaacttt atgaaccttt 1020
tccaagttga tctatccagt gacgtggcct ggtgggcgtt tcttcttgta cttatgtggt 1080
tttttggtt ttaatacaga cattttctc caaaaaaaaa aaaaaaaagg 1130

```

<210> 18

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL080235

<400> 18

ctttgaaaag caaggtttgt gctgcgcttc cagttccgaa aagcagatgt ttaagccctt 60

<210> 19

<211> 2498

<212> DNA

<213> Homo sapiens

<300>

<308> AL137540

<400> 19

```

gctgaaacga cagtcttgtc cctgtcagag aaatgacctg aacgaagagc ctcaacattt 60
tacacactat gcaatctatg atttcattgt caagggcagc tgcttctgca atggccacgc 120
tgatcaatgc atacctgttc atggcttcag acctgtcaag gccccaggaa cattccaca 180
ggtccatggg aagtgtatgt gtaagcacia cacagcaggc agccactgcc agcactgtgc 240
cccgttatac aatgaccggc catgggaggg agctgatggc aaaacggggg ctcccaacga 300

```

```

gtgcagaacc tgcaagtgtg atgggcatgc tgatacctgt cacttcgacg ttaatgtgtg 360
ggaggcatca ggggaatcgta gtggtggtgt ctgtgatgac tgtcagcaca acacagaagg 420
acagtattgc cagaggtgca agccaggctt ctatcgtgac ctgctggagac ccttctcagc 480
tccagatgct tgcaaaccgt gttcctgcca tccagtagga tcagctgtcc ttcctgccaa 540
ctcagtgacc ttctgcgacc ccagcaatgg tgactgccct tgcaagcctg ggggtggcagg 600
gcgacgttgt gacaggtgca tgggtgggata ctggggcttc ggagactatg gctgtcgacc 660
atgtgactgt gcggggagct gtgaccctat caccggagac tgcatacaga gccacacaga 720
catagactgg tatcatgaag ttcttgactt ccgtcccgtg cacaataaga gcgaaccaagc 780
ctgggagtgg gaggatgcgc aggggttttc tgcacttcta cactcaggta aatgcgaatg 840
taaggaacag acattaggaa atgccaaggc attctgtgga atgaaatatt catatgtgct 900
aaaaataaag attttatcag ctcatgataa aggtactcat gttgaggtca atgtgaagat 960
taaaaaggtc ttaaaatcta ccaaactgaa gattttccga ggaaagcgaa cattatatcc 1020
agaatcatgg acggacagag gatgcacttg tccaatcctc aatcctgggt tgggaatacct 1080
tgtagcagga catgaggata taagaacagg caaactaatt gtgaatatga aaagctttgt 1140
ccagcactgg aaaccttctc ttggaagaaa agtcatggat attttaaaaa gagagtgcag 1200
gtagcattaa gatggatagc acataatggc acttgtctat gtacaaaaca caaactttag 1260
agcaagaaga cctcagacag gaaactggaa ttttttaaag tgccaaaaca tatagaaatg 1320
tttgaatgca tgggtcttat ctaacttatc tcttctggac ccatgtttta atacagtttt 1380
atttcatgaa gagaaatgaa aacccttaca ctgatatctg ttttctatgg gactgattct 1440
gaaattctta actattaaga atattttaat agcagcatga catttagcag taatccatct 1500
agggcagtac ctctaacaag gacgccttcc agcttcagcg atgttactta cgtttgatgc 1560
tacttaaagt aatgaatgac gttttaagga atccctaacc ctactatcag aaaagggtgt 1620
tgttaaagag ccttctcttg tgtgttacgc atgaactttg gtctgtagggt gttaaatgga 1680
acctctccat gtgtatatag tatttccttg tataaagcac ttactacct accacttgtg 1740
ttgtgaacgt ttggtgactg ctgttgaaag aaggaaaagg gtgtgtgaga aagcctaactg 1800
aagcagcagc actgccacta catgtggaca aaagtgacca tataaaagaa gttgtgctat 1860
ttaactctga atacttgag aaactagggt aagatgcaac cagaaaggag aatatgtatg 1920
cgtgaagtct cagctttgag ctggaggcta gattccaaga tgacagccat gatgaaactt 1980
tttaaaaaac taaaccagaa gagactttta aataagagaa agaaatcata aatgtagaca 2040
tatgcttggc taaaggggaa atggacttta aattttaaag agctcatttg caatgcactt 2100
gtatacactt caaaaattat tgtagacaca gaatttgta tatttttgtg cttagtattt 2160
aaacctgaac attgaaacag ttttctcct tgtctttctt aacagtaata gtcattatat 2220
ttacctgttt tttaacacaa tgtatgtgat agtcaaaaaa tcacagtttt tcattattat 2280
tcactctctg taccacgca taaccactat acatagtttc ttttgtactt gaatatacaa 2340
aacatgaaca cagtgccata tgaataattt cacatacaga accttttttt ctctgaagtc 2400
ctgtggactt gcaaataat atatatattg ctttgttaat ttgtttttat atttcatata 2460
tgtaataaag gaatatgatc tgaaaaaaaa aaaaaaaa 2498

```

<210> 20

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL137540

<400> 20

```

tggaggctag attccaagat gacagccatg atgaaacttt ttaaaaaact aaaccagaag 60

```

<210> 21

<211> 914

<212> DNA

<213> Homo sapiens

<300>

<308> AL160131

<400> 21

```

cgcaccgcag gagcaacggt tggctcctgcg gctgtgatgt cgggtgttgag gcccttggac 60
aagctgcccc gcctgaaca cggccaccatc ttgctgggtg gcacggagga tgcctctctg 120
cagcagctgg cggactcgat gctcaaagag gactgcgcct ccgagctgaa ggtccacttg 180

```

```

gcaaagtccc tccctttgcc ctccagtgtg aatcggcccc gaattgacct gatcgtgttt 240
gtggttaatc ttcacagcaa atacagtctc cagaacacag aggagtccct gcgccatgtg 300
gatgccagct tcttcttggg gaaggtgtgt ttctctgccca caggtgctgg gcgggagagc 360
cactgcagca ttcaccggca caccgtggtg aagctggccc acacctatca aagccccctg 420
ctctactgtg acctggaggt ggaaggcttt agggccacca tggcgcagcg cctggtgcgc 480
gtgctgcaga tctgtgctgg ccacgtgccc ggtgtctcag ctctgaacct gctgtccctg 540
ctgagaagct ctgagggccc ctccctggag gacctgtgag ggtggctggc ccctgggctg 600
ccccttctca tggcttcgtg ctgactccat aaacattctc tgttgaggat gtccagtca 660
ggcttgacag gcccaggctc agcccgcgt ggctgggaag gttccctgca gtgccagtgc 720
tgacagaggg agagctgggc agaagcagcg agggggccca gctggcgaga ctgtagcccc 780
ctcccactcc cacactcact cttgcagagc ctgtgtcttt aagcagctgg cgtgttaca 840
ctccatttaa ggtttccttt gaacaaaagg tctgtggcta aaaaaagttt aaaaatcac 900
ggtctcattc acca 914

```

<210> 22
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL160131

<400> 22
 agctggcgtg ttacatctcc atttaaggtt tcctttgaac aaaaggtctg tggctaaaaa 60

<210> 23
 <211> 4753
 <212> DNA
 <213> Homo sapiens

<300>
 <308> D13642

<400> 23
 cttcaatcaa gtagccttcc cactgcagta cacacccagg aaatttgtca tccaccctga 60
 gtagtaacaac cttattatca ttgaaacgga ccacaatgcc tacactgagg ccacgaaagc 120
 tcagagaaaag cagcagatgg cagaggaaat ggtggaagca gcaggggagg atgagcggga 180
 gctggccgca gagatggcag cagcattcct caatgaaaac ctccctgaat ccattcttgg 240
 agctcccaag gctggcaatg ggcagtgggc ctctgtgatc cgagtgatga atccatttca 300
 agggaaacaca ctggaccttg tccagctgga acagaatgag gcagctttta gtgtggctgt 360
 gtgcagggtt tccaacactg gtgaagactg gtatgtgctg gtgggtgtgg ccaaggacct 420
 gatactaaac ccccgatctg tggcaggggg ctctgtctat acttacaagc ttgtgaacaa 480
 tggggaaaaa ctggagtttt tgcacaagac tctgttgaa gaggtccctg ctgctattgc 540
 cccattccag gggagggtgt tgattgggtg ggggaagctg ttgcgtgtct atgacctggg 600
 aaagaagaag ttactccgaa aatgtgagaa taagcatatt gccattata tctctgggat 660
 ccagactatt ggacataggg taattgtatc tgatgtccaa gaaagtttca tctgggttcg 720
 ctacaagcgt aatgaaaacc agcttatcat ctttgctgat gatacctacc cccgatgggt 780
 cactacagcc agcctcctgg actatgacac tgtggctggg gcagacaagt ttggcaacat 840
 atgtgtggtg aggctccac ctaacaccaa tgatgaagta gatgaggatc ctacaggaaa 900
 caaagccctg tgggaccgtg gcttgcctca tggggcctcc cagaaggcag aggtgatcat 960
 gaactaccat gtcggggaga cgtgtctgtc cttgcagaag accacgctga tccctggagg 1020
 ctcagaatca cttgtctata ccaccttgtc tggaggaatt ggcattcctt tgccattcac 1080
 gtcccattgag gaccatgact tcttccagca tgtggaatg cacctgcggg ctgaacatcc 1140
 cctctctgt gggcgggacc acctcagct tgcctctac tacttccctg tgaagaatgt 1200
 gattgatgga gacctctgtg agcagttcaa ttccatggaa cccaacaaac aaaagaacgt 1260
 ctctgaagaa ctggaccgaa ccccacccga agtgtccaag aaactcgagg atatccggac 1320
 ccgctacgcc ttctgagccc tcccttcccg gtggggcttg ccagagactg tgtgttttgt 1380
 ttccccacc accatcactg ccacctggct tctgccatgt ggcaggaggg tgactggata 1440
 attaagactg cattatgaaa gtcaacagct ctttccctc agctcttctc ctggaatgac 1500
 tggcttcccc tcaaattggc actgagattt gctacacttc tccccacctg gtacatgata 1560
 catgacccca ggttccagtg tagaacctga gtccccatt ccccaaagcc atccctgcat 1620

tgatatgtct	tgactctcct	gtctactttt	gcacacaccc	ttaattttta	attggttttc	1680
ttgtaaatac	agttttgtac	aatgttatct	ctgtgggagg	aaggaggcag	gctgtggtgg	1740
gactgggtag	ggtatagtat	cactcctgag	ttccactgct	ctagaatcta	accagaaata	1800
gaaacctagt	ttttaagggtg	actggcatcc	atgtgtcttg	ttctggagat	gaggatgtag	1860
gtgggaggtt	tgaacccaag	ttagagcagg	aagaactgag	tagactcctt	ccttccagat	1920
accgacttgg	acttgcgga	ctctgtggct	ccccaccccc	aggtctgtgg	tggtttcttt	1980
gttttttcct	ggttcttttt	gctgtgctga	tgaacatga	cctcaataac	catgtgtata	2040
cccacccctc	ttcccactgg	gtattgagga	agggtggctg	attcttcctc	ctcttctact	2100
ctgaggatgt	tagtatgggg	attttagcat	gaattccagc	tggggagtct	taacagatgc	2160
cccttttact	gatagagcac	ctaaagcgat	ctttggctcc	ataggacat	aggaagggtc	2220
agtacagaag	aacctagata	ctgccctgcc	cctgagaact	gtgtatatgt	ggggcctgtc	2280
tgcagcacc	atctcaggtg	ggttccagag	ggccttttagg	gtataatgag	agcctgttag	2340
gtggaagagg	cccagttcca	gaaatgttcc	agcccccccc	tgagaattcc	tcctgttttag	2400
ttgtgtggga	agccctcgtc	ttccaggctg	tccttgccgc	ttgaacctgg	agaagtgagc	2460
tcactgttct	caatacttca	caaatgtaaa	actttctttc	gtctgcatgt	gctcagcat	2520
ctaaattgag	caaatgatct	ggtgagcact	gggttagaat	caggaatggg	ggaatacaat	2580
ctgaacctct	cagagcccag	aacagagggg	tcctgacact	gtgacactgt	ctcctggaac	2640
taagtattct	ttgaatcatg	acttggtttt	agatcagtca	agagagaccc	aggttttgcc	2700
aggaatcgaa	tccttaaata	acatgttttt	ttctcactta	gctcatgaat	ttgcatagta	2760
gacagtagtt	ctgaattaga	ttttgaaaac	ctaatttcag	ggctcatttt	ttctgtggc	2820
cctaaatcca	ttctatcaaa	ttgtgtgata	ctgacatgca	gtcatctgag	gaactcagcg	2880
tagatacttg	agcagctcct	cgcctctttt	ctaactcaag	tttgactaaa	atacatacac	2940
tcctgtacaga	aggtaggggg	ttatgtaaaga	aaggaaaacc	taatctatgg	aatcaggagt	3000
tgtcaccacc	gagcttcctc	tggaaagtctg	cccatcagct	tgcttgttct	ctgttaagag	3060
gaagggctag	gacaaggatt	tgggcttgaa	tatgtggaaa	ggaattttca	tagttgttgc	3120
tgcaggacct	acaaaagttt	aaaatttagat	tggatgtgac	tcaatgacaa	gtcccatctg	3180
tgtaattgtt	aaggggacct	gattgactcc	tgtggtttga	ttgagcaacc	aggtaaaatg	3240
agacctctct	ccagcttttg	caaaacccat	cagaggctgc	tgcagaactc	agacagaggg	3300
atctgcccct	gggtttgtct	ccatcctgtt	ccattgctaa	gcccttgtga	cttggatcct	3360
aggactgaaa	agtttttagc	tgcctcagct	ttcccctgac	cttactggca	gaggttctgc	3420
agatgtttcc	tttggaagat	ctcttgccaa	gaatagcatt	cctttggagg	aggggggttc	3480
tagttggaat	gttgcttttc	ttggttagtg	taaattgtatt	gctagtgaga	cagctgcagg	3540
cgctggaaaa	ggctcgtctc	acagggagag	tgctggctcc	cagaatgtgt	gctgttccca	3600
cgctgctgcc	tttcttgagc	ttgttagagg	aaagccagaa	aggcattcag	atgggtaacg	3660
tctggctttt	aaattttttt	taattcctaa	gttctgtttt	attttttaat	tttttaaaaa	3720
aaatttttatt	agagacagtc	tctctctctt	gcctagctgg	gagtgcagtg	gagtgatcat	3780
agctcactga	ggcttgaact	cctgggctcg	agcaatccac	ctcagcctcc	agagtagggg	3840
agactacaga	tgtgtgccac	catactcagc	tagtttttaa	actttcgtag	agacagggtc	3900
tcctgtgtgt	gccaggtctg	gcctcgaact	cctgacctca	aaaaatcttc	ctgccttggc	3960
ctcccagcgc	tttgagaggc	tgaggcagga	ggatcccttg	agcccaggag	tttgagacca	4020
gcctgggcaa	catgacaaaa	ccccatctct	caaaaaatac	aaaaattggc	caggcatggt	4080
gggtgcacact	tgtagtccca	gtaattaggg	ggctgagaca	ggaggatcac	ttcagcctat	4140
gagtttgagg	ctgcagtgag	ctgtgattgc	gccactacac	tccagcctgg	atgacaggac	4200
gaaacctgtc	tcaaaaaacac	caaaaaacaa	aaaccggctc	cctgggggtca	tggttagcaca	4260
aacgcacatg	actgagtgtc	caggggttct	gaggcttgtc	cgctgacctg	gggctctggc	4320
cctggggagat	ctgggggacc	tgctgtccta	tatgtgatgc	tttgaaagaa	aggggcatca	4380
ttccaagcca	agaggcccca	gagagggcac	cgtgggggtg	tcaggcctct	gtgaggcccc	4440
agtgagatcc	tgtggctgtg	cccccatcac	ctccaccac	tctgcccctc	cactagctgc	4500
ccaacggatg	aatcaacgcc	ttggcagagt	tttccagcag	ggccttgcag	agagtgtgtg	4560
tgacctgtgt	ggccactgcc	ttggggacgg	gtgaggagtt	agcctggaac	attccagcgt	4620
gggcattatt	gtcctgttgc	aagttcaggg	caaaaccagg	aatccagttt	tgtcgatcca	4680
attgagaaaa	catttcatga	acaactactt	gtggcatgca	ttggcactcg	gaataaagcg	4740
cactattgtc	act	4753				

<210> 24

<211> 60

<212> DNA

<213> Homo sapiens

<300> .

<308> D13642

<400> 24
aaaccaggaa tccagttttg tccgatccaat tgagaaaaca tttcatgaac aactacttgt 60

<210> 25
<211> 2591
<212> DNA
<213> Homo sapiens

<300>
<308> D25328

<400> 25
cccggacgtg cggctcccc t cggcctcctc gccatggacg cggacgactc cggggc cccc 60
aagggctcct tgcggaagt t cctggagcac ctctccgggg cgggcaaggc catcgg cgtg 120
ctgaccagcg gcggggatgc tcaaggatat aacgctgccg tccgtgccgt ggtgcg catg 180
ggtatctacg tgggggcca a ggtgtacttc atctacgagg gctaccaggg catggt ggac 240
ggaggctcaa acatcgca g ggcgactgg gagagtgtct ccagcatcct gcaagt gggc 300
gggacgatca ttggcagtgc gcggtgccag gccttcgcga cgcgggaagg ccgcct gaag 360
gctgcttgca acctgctgca gcgcggcatc accaacctgt gtgtgatcgg cgggga cggg 420
agcctcaccg gggccaacc t ctcccggaag gagtggagtg ggctgctgga ggagct gggc 480
aggaacggcc agatcgataa ggaggcgtg cagaagtacg cctacctcaa cgtggt gggc 540
atggtgggct ccatcgaca a tgattttctgc ggcaccgaca tgaccatcgg cacgga ctcc 600
gccctgcaca ggatcatcga ggtcgtcgac gccatcatga ccacggcca gagcca ccag 660
aggaccttcg ttctggagg t gatgggacga cactgtgggt acctggccct ggtgag tgcc 720
ttggcctgcg gtgcggact g ggtgttcctt ccagaatctc caccagagga aggtg ggag 780
gagcagatgt gtgtcaaac t ctcgagaaac cgtgcccgga aaaaaaggct gaatat tatt 840
attgtggctg aaggagcaa t tgatacccaa aataaaccca tcacctctga gaaaat caaa 900
gagcttgctg tcacgcagc t gggctatgac acacgtgtga ccacctcgg gcacgt gcag 960
agaggaggga ccccttcggc attcgacagg atcttggcca gccgcatggg agtggaggca 1020
gtcatcgcct tgctagaggc caccocggac accccagctt gcgtcgtgtc actgaa cggg 1080
aaccacgccc tgccctgc c gctgatggag tgcgtgcaga tgactcagga tgtgca gaag 1140
gcgatggacg agaggagat t tcaagatgcg gttcgactcc gagggaggag ctttgc gggc 1200
aacctgaaca cctacaagc g acttgccatc aagctgccgg gcacccgcgg ctgggatgaa cgcggc cgta 1260
aattgcaacg tagctgtca t caacgtgggg gccacccgcgg ctgggatgaa cgcggc cgta 1320

cgctcagctg tgcgcgtggg cattgccgac ggccacagga tgctcgccat ctatga tggc 1380
tttgacggct tcgccaagg ccagatcaaa gaaatcggct ggacagatgt cggggg cttg 1440
accggccaag gaggtccat tcttgggaca aaacgcgttc tcccggggaa gtactt ggaa 1500
gagatcgcca cacagatgcg cacgcacagc atcaacgcgc tgctgatcat cgggtg gattc 1560
gaggcctacc tgggactcct ggagctgtca gccgcccggg agaagcacga ggagt tctgt 1620
gtccccatgg tcatggttcc cgtactgtg tccaacaatg tgccgggttc cgattt cagc 1680
atcggggcag acacgcctc gaacactatc accgacacct gcgaccgat caagc agtcc 1740
gccagcggaa ccaagcggc cgtgttcac atcgagacca tgggcccgtc ctgtg gctac 1800
ctggccaaca tgggggggct cgcggccgga gctgatgccg catacatttt cgaag agccc 1860
ttcgacatca gggatctgca gtccaacgtg gagcacctga cggagaaaat gaaga ccacc 1920
atccagagag gccttgtgct cagaaatgag agctgcagtg aaaactacac caccga cttc 1980
atttaccagc tgtattcaga agagggcaaa ggcgtgtttg actgcaggaa gaacgt gctg 2040
ggtcacatgc agcagggtgg ggcaccctct ccatttgata gaaacttttg aaccaaaatc 2100
tctgccagag ctatggagt g gatcactgca aaactcaagg agggccgggg cagagg aaaa 2160
aaatttacca ccgatgatc catttgtgtg ctgggaataa gcaaaagaaa cgttat tttt 2220
caacctgtgg cagagctgaa gaagcaaacg gattttgagc acaggattcc caaaga acag 2280
tggtggctca agctacggc cctcatgaaa atcctggcca agtacaaggc cagctatgac 2340
gtgtcggact caggccagc t ggaacatgtg cagecctgga gtgtctgacc cagtc ccgc 2400
tgcatgtgcc tgcagccacc gtggactgtc tgtttttgta acacttaagt tatttt atca 2460
gcactttatg cacgtattat tgacattaat acctaatcgg cgagtggcca tctgc cccac 2520
cagctccagt gcgtgctgtc tgtggagtgt gtctcatgct ttcagatgtg catat gagca 2580
gaattaatta a 2591

<210> 26
<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> D25328

<400> 26
tattttatca gcactttatg cacgtattat tgacattaat acctaatacgg cgagtgccca 60

<210> 27
<211> 2573
<212> DNA
<213> Homo sapiens

<300>
<308> D50402

<400> 27
gaatcgcccg atgtgaaccg aatgttgatg taagaggcag ggcactcggc tgcgga tggg 60
taacagggcg tgggctggca cacttacttg caccagtgcc cagagagggg gtgcagcctg 120
aggagctgcc cagagcaccg ctcacactcc cagagtacct gaagtcggca tttcaa tgac 180
aggtgacaag ggtcccca aa ggctaagcgg gtccagctat ggttccatct ccagcccgac 240
cagcccgacc agcccagggc cacggcaagc acctcccaga gagacctacc tgagtgaaga 300
gatccccatc ccagacacaa aaccgggcac cttcagcctg cggaagctat gggcct tcac 360
ggggcctggc ttctctcatga gcattgcttt cctggaccca ggaaacatcg agtcagatct 420
tcaggctggc gccgtggcgg gattcaaaact tctctgggtg ctgctctggg ccaccgtgtt 480
gggcttgctc tgccagcgac tggctgcacg tctgggcgtg gtgacaggca aggact tggg 540
cgaggctctgc catctcta ct accctaaggt gccccgcacc gtctctctggc tgacca tcga 600
gctagccatt gtgggctc cg acatgcagga agtcatcggc acggccattg cattca atct 660
gctctcagct ggacgaat cc cactctgggg tggcgctctc atcaccatcg tggaca cctt 720
cttcttctctc ttctctcga ta actacgggct gcggaagctg gaagcttttt ttggac tctt 780
tataaccatt atggccttga cctttggcta tgagtatgtg gtggcgcgctc ctgagc aggg 840
agcgcttctt cggggcctgt tcctgcctc gtgcccgggc tgcggccacc ccgagc tgct 900
gcaggcggtg ggcattgt tg gcgccatcat catgcccac aacatctacc tgact cggc 960
cctgggtcaag tctcgaga ga tagaccgggc ccgccgagcg gacatcagag aagcca acat 1020
gtacttctctg attgaggc ca ccatcgccct gtccgtctcc tttatcatca acctct ttgt 1080
catggctgtc tttgggca gg ccttctacca gaaaaccaac caggctgcgt tcaaca tctg 1140
tgccaacagc agcctcca cg actacgccaa gatcttcccc atgaacaacg ccaccgtggc 1200
cgtggacatt taccagggg gcgtgatcct gggctgcctg ttgggccccg cggccc tcta 1260
catctggggc ataggtct cc tggcggtctg gcagagctcc accatgacgg gcacct acgc 1320
gggacagttc gtgatgga gg gcttctctgag gctgcgttgg tcacgcttcg ccctgtctct 1380
cctcaccgc tcctgcgc ca tcctgcccac cgtgctcgtg gctgtcttcc gggacc tgag 1440
ggacttgctg ggctccta tg atctgctcaa cgtgctgcag agcctgctgc tcccg tctgc 1500
cgtgctgccc atctcacc gt tcaccagcat gccaccctc atgcaggagt ttgcca atgg 1560
cctgctgaac aaggctgt ca cctcttccat catggtgcta gtctgcgcca tcaacctcta 1620
cttcgtggtc agctatct gc ccagcctgcc ccaccctgcc tacttcggcc ttgcagcctt 1680
gctggccgca gcctacct gg gcctcagcac ctacctggtc tggacctgtt gccttgccca 1740
cggagccacc tttctggccc acagctccca ccaccacttc ctgtatgggc tccttgaaga 1800
ggaccagaaa ggggagac ct ctggctaggc ccacaccagg gcctggctgg gaggggcatg 1860
tatgacgtga ctggcctgct ggatgtggag ggggcgcgtg caggcagcag gatggagtgg 1920
gacagttcct gagaccagcc aacctggggg ctttagggac ctgctgtttc ctagcgcagc 1980
catgtgatta ccctctgggt ctcaagtgtc tcatctgtaa aatggagacg ccaccaccct 2040
tgccatggag gtttaagca ct ttaacacagt gtctggcact tgggacaaaa acaaa caaac 2100
aaacaaaaaa catttcaaaa ggtatttatt gagcacctgc aggcgtgacc tgacagccca 2160
aggggtgggtg ggggtgagggc ttgaggactt gggcgggaca caggctccaa actggagctt 2220
gaaatagtgt ctgatgaa tg ttaaattatc tatctatcta tttatttatt tatttgagac 2280
agggaaaggg tctccctctg ttgccaaggc tggagtgcag tggcgcaatc ttaactcatt 2340
gcaacctcca ccttctgggt tcaagcgatt ctctttatc agccccggga gtggcgcgcg 2400
ccaccagcc cagctaat tt gtgtattttc agcagagacg gggtttgcca tgctgcccag 2460
gctggctctg aactgctgga ttcaagtgat ccgccatct ccgtctccca aagtgcctggg 2520

aattacaggc gtgagc cacc aaaacccggc ctgattaaag ttaaataaat acg 2573

<210> 28

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 28

tgagggttaa gcactttaac acagtgtctg gcacttggga caaaaacaaa caaa caaaca 60

<210> 29

<211> 3672

<212> DNA

<213> Homo sapiens

<300>

<308> L27560

<400> 29

acatgtgcat	atttcattcc	ccaggcagac	atTTTTtaga	aatcaataca	tgccccaata	60
ttggaaagac	ttgttcttcc	acggtgacta	cagtacatgc	tgaagcgtgc	cgtttcagcc	120
ctcatttaat	tcaatttgta	agtagcgcac	gagcctctgt	gggggaggat	aggctgaaaa	180
aaaaaagtgg	gctcgtatTT	atctacagga	ctccatatag	tcatatatag	gcataataat	240
ctatgctttt	tctttgtttt	tttctttctt	cctttctttc	aaagggtttgc	attaactttt	300
caaagtagtt	cctatagggg	cattgaggag	cttcctcatt	ctgggaaaac	tgagaaaacc	360
catattctcc	taatacaacc	cgtaatagca	tttttgcttg	cctcgaggca	gagt tccccg	420
tgagcaataa	actcagcttt	tttgtggggc	acagtactgg	atTTgacagt	gatt ccccac	480
gtgtgttcat	ctgcacccac	cgagccaggc	agaggccagc	cctccgtggt	gcacacagca	540
cgcgcctcag	tccatcccat	tttagtcttt	aaacctcag	gaagtcacag	tctc cggaca	600
ccacaccaca	ttgagcccaa	cagggtccacg	atggatccac	ctagtccac	cccagccttt	660
ttctttcatc	tgaacagaat	gtgcattttt	ggaagcctcc	ctcactctcc	atgc tggcag	720
agcaggaggg	agactgaagt	aagagatggc	agaggagat	ggtggcaaaa	aggt ttagat	780
gcaggagaac	agtaagatgg	atggttccgg	ccagagtcga	tgtggggagg	aacagagggc	840
tgaagggaga	gggggctgac	tgttccattc	tagctttggc	acaaagcagc	agaaaggggg	900
aaaagccaat	agaaa tttcc	ttagcttccc	caccatatgt	atTTtcatgg	atttcagagg	960
aaagagagga	aaatggggga	atgggttgca	aaatagaaat	gagcttaatc	caggcccgag	1020
agccagggaa	ggtgagtaac	cttaggaggg	tgctagactt	tagaagccag	atagggaagaa	1080
tcagtctaaa	ctggccatgc	tttggaaagg	acaagactat	gtgctccgct	gccaccttc	1140
agcctgcaat	gagggactga	ggcccacgag	tctttccagc	tcttctcca	ttctggccag	1200
tccctgcatc	ctccc tgggg	tggaggatcg	aaggaaagct	gggacaagca	gggaacgcac	1260
gattcaggga	tgctgtcact	cggcagccag	attccgaaac	tccattctc	caatgacttc	1320
ctcaaccaat	gggtggcctt	gtgactgttc	tttaaggctg	aagatatcca	ggaaaggggg	1380
cttggacact	ggccaaggag	accccttcgt	gctgtggaca	cagctctctt	cactctttgc	1440
tcatggcatg	acacagcgga	gaccgcctcc	aacaacgaat	ttggggctac	gaagaggaat	1500
agcgaaaaag	caaactgtgt	tcaactgatg	ggaaccctat	agctatagaa	cttgggggct	1560
atctcctatg	cccctggaca	ggacagttgg	ctggggacag	gagaagtgtc	caatcttcat	1620
gagacaaagg	ggccc gatca	aggcagccac	aaggccttga	cctgccgagt	cagcatgccc	1680
catctctctc	gacagctgtc	ccctaaaccc	aactcacgtt	tctgtatgtc	ttaggccagt	1740
atcccaaacc	tcttc cagct	cactgttctt	tccaccatt	ctccctttgc	atcttgagca	1800
gttatccaac	tagga tctgc	caagtggata	ctgggggtgc	actcccctga	gaaaagactg	1860
agccaggaac	tacaagctcc	ccccacattc	ctcccagcct	ggacctaat	cttgagaggg	1920
gctctctctt	cacggactgt	gtctggactt	tgagcaggct	tctgccctt	gcgttggttc	1980
tttgtgcca	gccat caggt	gggggattag	agcctggtgt	aagtgcgcca	gactcttccg	2040
gtttccaaag	ttcgt gcctg	cgaacccaaa	cctgtgagtc	tcttctgcat	gcaggagttt	2100
ctcctgggca	gctggtcact	ccccagagaa	gctgggcctt	catggacaca	tggaaactaag	2160
cctcccaa	gggagtctctg	gctgagccca	gggtggggag	atcctgggaa	gggaggcact	2220
ggaggaagac	ggcac ctctt	cccccatggc	aggggtgtgag	ggaggcaggt	ttggaaatggt	2280
gcgagtatgg	caatc taagc	aggggtctgg	tctctttgac	tccaggctcg	ctt tggccga	2340

ctgtctgctc	accagagac	cttggactcc	ggactatcca	tggtccgaa	tctaagtgct	2400
gccactccc	atgtcacac	ccacagaagg	tcttcccatc	cccttagat	tcgtgcctca	2460
ctccaccagt	gaggaagatg	cctctgtctt	tcccacgact	gccaggagat	agggaagccc	2520
agccaggact	gacctcctt	cctccagcct	gccctgaccc	acctggcaaa	gcagggcaca	2580
tggggaggaa	gagactggaa	cctttctttg	acagccaggc	ctagacagac	aggcctgggg	2640
acactggccc	atgaggggag	gaaggcaggc	gcacgaggtc	cagggaggcc	cttttctgat	2700
catgcccctt	ctctcccacc	ccatctcccc	accaccacct	ctgtggcctc	catggtacct	2760
ccacagggtc	ggcctccct	agagggtggg	cctcaaccac	ctcgtcccg	cacgcaccgg	2820
ttagtgagac	agggctgcca	cgcaaccgcc	aagccccct	caagggtgga	cagtaccccg	2880
gacccatcca	ctcactcctg	agagggtccg	gcccagaatg	ggaacctcag	agaagagctc	2940
taaggagaag	aaacccata	gcgtcagaga	ggatatgtct	ggcttccaag	agaaaggagg	3000
ctccgttttg	caaagtggag	gagggacgag	ggacaggggt	ttcaccagcc	agcaacctgg	3060
gccttgtact	gtctgtgttt	ttaaaaccac	taaagtgcaa	gaattacatt	gcactgtttc	3120
tccacttttt	attttctctt	aggcttttgt	ttctatttca	aacatacttt	cttgggtttc	3180
taatggagta	tatagttag	tcatttcaca	gactctggcc	tcctctcctg	aaatcctttt	3240
ggatggggaa	agggaaagg	gggagggtcc	gaggggaagg	ggacccagc	ttccctgtgc	3300
ccgctcacc	cactccacca	gtccccggtc	gccagccgga	gtctcctctc	taccgccact	3360
gtcacaccgt	agcccatg	gatagcacag	ttgtcagaca	agattccttc	agattccgag	3420
ttgctaccgg	ttgttttcgt	tgttgttgtt	gttgtttttc	tttttctttt	tttttttgaa	3480
gacagcaata	accacagtac	atattactgt	agttctctat	agttttacat	acattcatac	3540
cataactctg	ttctctctc	ttttttgttt	tcaactttaa	aaacaaaaat	aaacgatgat	3600
aatctttact	ggtgaaagg	atggaaaaat	aatcaacaa	atgcaaccag	tttgtgagaa	3660
aaaaaaaaa	aa	3672				

<210> 30

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> L27560

<400> 30

agcaacctgg gccttgtact gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt 60

<210> 31

<211> 1416

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base

<222> 1 ... 1416

<223> n = a,c,g, or t

<300>

<308> M55914

<400> 31

aggaattccg	gaattccgga	attccgatgg	atggaacaga	aaataaatct	aagtttggtg	60
cgaacgccat	tctgggggtg	tcccttgccg	tctgcaaagc	tggtgccgtt	gagaaggggg	120
tcccctgtac	cgccacatcg	cgtacttgge	tggcaacttc	gaagtcatcc	tgccagtccc	180
ggcgttcaag	tgtcatcatc	aatggcggtt	ctcatgctgg	caacaagctg	gccatgcaga	240
gtctgtcctc	ccagtcgggtg	cagcaaactc	agggaaagcca	tgccgcattg	gagcagaggt	300
ttaccacaac	ctgaagaatg	tcatcaagga	gaaatatggg	aaagatgcca	ccaatgtggg	360
gatttgccgcg	ggtttgctcc	caacatcctg	gagaataaag	aaggcctgga	gctgctgaag	420
actgctattg	gaaagcctgg	cctacactgt	aaagggtggtc	atggcatgga	cgtagcggcc	480
tccgagttct	tcaggtcagg	gaactatgac	ctggacttca	agtctcccga	tgacccagc	540
aggtacatct	cgcctgacca	gctggctgac	ctgtacaagt	ccttcatcaa	ggactacca	600
gtggtgtcta	tcgaagatcc	ctttgaccag	gatgactggg	gagcttcaga	agttcacagc	660
cagtgcagga	atccaggtag	tggggggatg	actcacagtg	accaacccaa	agaggatcgc	720

```

caaggcgtga acgagaagtc ctgcaactgc ctctctgtca aagtcaacca gattggctcc 780
gtgaccgagt ctcttcaggc gtgcaagctg gccagggcca atgggtgggg cgtc atgggtg 840
tctcatcggt cgggggagac tgaagatacc ttcacgctg acctgggtgt ggggctgtgc 900
actggggcag atcaagactg gtgccccttg ccgatcacgc gcttggccaa gtacaaccag 960
ctcctcagaa ttgaagagga gctgggcagc aaggctaagt ttgccggcag gaacttcaga 1020
aacccttggt ccaagtaagc tgtgggcagg caagccttcg gtcacctgtt ggctacagac 1080
ccctcccttg gtgtcagctc aggcagctcg agggccccga ccaacacttg cagggggtccc 1140
tgctagttag cgcccaccgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagctccctg gaagccctgt tggcagctct agctttgcag ttgtgtaatt ggcccaagtc 1260
attgtttttc tcgcttact ttccaccaag tgtctagagt catgtgagcc tngtgtcatc 1320
tccgggggtg ccacaggcta gatccccggg ggttttgtgc taaaataaa aagctcagc 1380
gacccatgaa aaaaaaaaag gaattccgga attccg 1416

```

<210> 32
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M55914

```

<400> 32
gtaccgcttc cttagaactc tacagaagcc aagctccctg gaagccctgt tggcagctct 60

```

<210> 33
 <211> 2517
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M96577

```

<400> 33
ggaattccgt ggccgggact ttgcaggcag cggcggccgg gggcggagcg ggatcgagcc 60
ctcgccgagg cctgc cgcca tgggcccggc ccgcccgcgc cgctgtcac ccggggccgcg 120
cgggcccgtga gcgtcatggc cttggccggg gccctgctgg gcggcccatg cgcgcccggcg 180
ctggaggccc tgctc ggggc cggcgcgctg cggctgtctg actcctcgca gatcgtcac 240
atctccgccc cgcaggacgc cagcgccccg ccggctccca ccggccccgc ggcggcccgc 300
gccggccccct gcgac cctga cctgctgctc ttgccacac cgcaggcgcc ccggcccaca 360
cccagtgcgc cgcggccccgc gctcggccgc ccgcccgtga agcggagggt ggaacctggaa 420
actgaccatc agtac ctggc cgagagcagt gggccagctc ggggcagagg ccgcacatcca 480
ggaaaagggtg tgaaa tcccc ggggggagaag tcacgctatg agacctact gaactctgacc 540
accaagcgct tcctggagct gctgagccac tcggctgacg gtgtcgtcga cctgaactgg 600
gctgccgagg tgctgaagggt gcagaagcgg cgcactatg acatcaccaa cgtccttgag 660
ggcatccagc tcatt gccaa gaagtccaag aaccacatcc agtggctggg cagccacacc 720
acagtgggcg tcggc ggacg gcttgagggg ttgaccagg acctccgaca gctgcaggag 780
agcgagcagc agctggacca cctgatgaat atctgtacta cgcagctgcg cctgctctcc 840
gaggacactg acagc cagcg cctggcctac gtgacgtgtc aggaccttcg tagcattgca 900
gaccctgcag agcagatggt tatggtgatc aaagcccctc ctgagacca gctccaagcc 960
gtggactctt cggagraactt tcagatctcc cttaagagca aacaaggccc gatcgatgtt 1020
ttcctgtgcc ctgaggagac cgtaggtggg atcagccctg ggaagacccc atcccaggag 1080
gtcacttctg aggaggagaa cagggccact gactctgcca ccatagtgtc accaccacca 1140
tcactctccc cctcatccct caccacagat ccagccagt ctctactcag cctggagcaa 1200
gaaccgctgt tgtcc cggat gggcagcctg cgggctcccg tggacgagga ccgcctgtcc 1260
ccgctggtgg cggcc gactc gctcctggag catgtgcggg aggacttctc cggcctcctc 1320
cctgaggagt tcatcagcct ttccccaccc cagcaggccc tcgactacca cttcgggcctc 1380
gaggagggcg agggcatcag agacctcttc gactgtgact ttggggacct cacccccctg 1440
gattttctgac agggc ttgga gggaccaggg tttccagagt agctcacctt gtctctgcag 1500
ccctggagcc cctgtccct ggccgtcctc ccagcctgtt tggaaacatt taatttatat 1560
ccctctctc tgtct ccaga agcttctagc tctggggctc ggctaccgct aggaggctga 1620

```

```

gcaagccagg aaaggaagga gtctgtgtgg tgtgtatgtg catgcagcct acacccacac 1680
gtgtgtaccg ggggtgaatg tgtgtgagca tgtgtgtgtg catgtaccgg ggaatgaagg 1740
tgaacataca cctctgtgtg tgcactgcag acacgcccc a gtgtgtccac atgtgtgtgc 1800
atgagtcctat ctctgctgct gggggggctc taactgcact ttcggccctt ttgctcgtgg 1860
ggtcccacaa ggcccagggc agtgccctgct cccagaatct ggtgctctga ccaggccagg 1920
tggggagggt ttggctgggt gggcgtgtag gacgggtgaga gcacttctgt cttaaagggt 1980
ttttctgatt gaagctttta tggagcgtaa tttatttatc gaggcctctt tggtagacct 2040
ggggaatcag caaaagggga ggaggggtgt ggggttgata cccaactcc ctctaccctt 2100
gagcaagggc aggggtccct gagctgttct tctgccccat actgaaggaa ctgaggcctg 2160
ggtgatttat ttattgggaa agtgaggag ggagacagac tgactgacag ccatgggtgg 2220
tcagatgggt gggtggggcc tctccagggg gccagttcag ggcccagctg ccccccagga 2280
tggatatgag atgggagagg tgagtggggg accttcaact atgtgggcag gaggggtggg 2340
gaaggcctcc ccagcccag accctgtggg cctcctgca gtgtctgaag cgcctgcctc 2400
cccactgctc tgccccaccc tccaatctgc actttgattt gcttctaac agctctgttc 2460
cctcctgctt tggttttaat aatatatttg atgacgttaa aaaaaggaat tcgatat 2517

```

<210> 34

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> M96577

<400> 34

```

gtaggacggg gagagcactt ctgtcttaaa ggttttttct gattgaagct ttaatggagc 60

```

<210> 35

<211> 4437

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000057

<400> 35

```

gcgcggcgcc cgtgggttgc gcgcgggaag tttggatcct ggttccgtcc gctaggagtc 60
tgcgtgcgag gattatggct gctgttcctc aaaataatct acaggagcaa ctagaacgtc 120
actcagccag aacacttaat aataaattaa gtctttcaaa accaaaattt tcagggtttca 180
cttttaaaaa gaaaacatct tcagataaca atgtatctgt aactaatgtg tcagtagcaa 240
aaacacctgt attaagaaat aaagatgtta atgttaccga agacttttcc ttcagtgaac 300
ctctacccaa caccacaaat cagcaaaggg tcaaggactt ctttaaaaat gctccagcag 360
gacaggaaac acagagaggt ggatcaaaat cattattgcc agatttcttg cagactccga 420
aggaagtgtg atgcactacc caaaacacac caactgtaaa gaaatcccgg gatactgctc 480
tcaagaaatt agaatttagt tcttcaccag attctttaag taccatcaat gattgggatg 540
atatggatga ctttgatact tctgagactt caaaatcatt tgttacacca ccccaaagtc 600
actttgtaag agtaagcact gctcagaaat caaaaaggg taagagaaac ttttttaaag 660
cacagcttta tacaacaaac acagtaaaga ctgatttgcc tccacctcc tctgaaagcg 720
agcaaataga ttgtactgag gaacagaagg atgactcaga atggttaagc agcgatgtga 780
tttgcatcga tgratggccc attgctgaag tgcataataa tgaagatgct caggaaagtg 840
actctctgaa aactcatttg gaagatgaaa gagataatag cgaaaagaag aagaatttgg 900
aagaagctga attacattca actgagaaag ttccatgtat tgaatttgat gatgatgatt 960
atgatacgga tttgttcca ccttctccag aagaaattat ttctgcttct tcttctctct 1020
caaaatgcct tagtacgtta aaggaccttg acacatctga cagaaaagag gatgttctta 1080
gcacatcaaa agatcttttg tcaaaacctg agaaaatgag tatgcaggag ctgaatccag 1140
aaaccagcac agactgtgac gctagacaga taagtttaca gcagcagctt attcatgtga 1200
tggagcacat ctgtaaatta attgatacta ttcctgatga taaactgaaa cttttgatt 1260
gtgggaacga actgcttcag cagcggaaca taagaaggaa acttctaacg gaagtagatt 1320
ttaataaaaag tgatgccagt cttcttggct cattgtggag atacaggcct gatttacttg 1380
atggccctat ggaggggtgat tcttgcctta cagggaattc tatgaaggag ttaaatTTTT 1440
cacaccttcc ctcaaattct gtttctcctg gggactgttt actgactacc acctaggaa 1500

```

agacaggatt	ctctgccacc	aggaagaatc	tttttgaaag	gcctttattc	aatacccatt	1560
tacagaagtc	ctttgtaagt	agcaactggg	ctgaaacacc	aagactagga	aaaaaaaaatg	1620
aaagctctta	tttccagga	aatgttctca	caagcactgc	tgtgaaagat	cagaataaac	1680
atactgcttc	aataaatgac	ttagaaagag	aaacccaacc	ttcctatgat	attgataatt	1740
ttgacataga	tgactttgat	gatgatgatg	actgggaaga	cataatgcat	aatttagcag	1800
ccagcaaate	ttccacagct	gcctatcaac	ccatcaagga	aggtcggcca	attaaatcag	1860
tatcagaaag	actttctca	gccaagacag	actgtcttcc	agtgtcatct	actgctcaaa	1920
atataaactt	ctcagagtca	attcagaatt	atactgacaa	gtcagcacaa	aatttagcat	1980
ccagaaatct	gaaacatgag	cgtttccaaa	gtcttagttt	tcctcataca	aaggaaatga	2040
tgaagatttt	tcataaaaaa	tttggcctgc	ataatttttag	aactaatcag	ctagaggcga	2100
tcaatgctgc	actgcttggg	gaagactggt	ttatctctgat	gccgactgga	ggtggtgaaga	2160
gtttgtgtta	ccagctccct	gcctgtgttt	ctcctggggg	cactgttgtc	atttctccct	2220
tgagatcact	tatcgtagat	caagtcctaa	agctgacttc	cttggatatt	ccagctacat	2280
atctgacagg	tgataagact	gactcagaag	ctacaaatat	ttacctccag	ttatcaaaaa	2340
aagacccaat	cataaaactt	ctatatgtca	ctccagaaaa	gatctgtgca	agtaacagac	2400
tcattttctac	cttggagaat	ctctatgaga	ggaagctctt	ggcacgtttt	gttattgatg	2460
aagcacattg	tgtcagtcag	tggggacatg	atcttctgtca	agattacaaa	agaatgaata	2520
tgcttcgcca	gaagtctcct	tctgttccgg	tgatggctct	tacggccaca	gctaatocca	2580
gggtacagaa	ggacaacctg	actcagctga	agattctcag	acctcaggtg	tttagcatga	2640
gctttaacag	acataatctg	aaatactatg	tattaccgaa	aaagcctaaa	aaggtggcat	2700
ttgattgcct	agaatggatc	agaaagcacc	acccatatga	ttcagggata	atttactgcc	2760
tctccaggcg	agaatgtgac	accatggctg	acacgttaca	gagagatggg	ctcgctgctc	2820
ttgcttacca	tgctggcctc	agtgattctg	ccagagatga	agtgcagcag	aagtggatta	2880
atcaggatgg	ctgtcagggt	atctgtgcta	caattgcatt	tggaatgggg	attgacaaac	2940
cggacgtgcg	atctgtgatt	catgcacttc	tcctaaaatc	tgtggagggt	tactaccaag	3000
aatctggcag	agctggaaga	gatggggaaa	tatctcaactg	cctgcttttc	tatacctatc	3060
atgatgtgac	cagactgaaa	agacttataa	tgatggaaaa	agatggaaac	catcatacaa	3120
gagaaactca	cttcaataat	ttgtatagca	tggtacatta	ctgtgaaaat	ataacggaat	3180
gcaggagaat	acagcttttg	gcctactttg	gtgaaaatgg	atttaatcct	gattttttgta	3240
agaaacaccc	agatgtttct	tgtgataatt	gctgtaaaac	aaaggattat	aaaacaagag	3300
atgtgactga	cgatgtgaaa	agtattgtaa	gattttgttca	agaacatagt	tcatacacaag	3360
gaatgagaaa	tataaaacat	gtaggtcctt	ctggaagatt	tactatgaat	atgctggctg	3420
acattttctt	ggggagtaag	agtgcacaaa	tccagtcagg	tatatttgga	aaaggatctg	3480
cttattctacg	acacaatgcc	gaaagacttt	ttaaaaagct	gatacttgac	aagatttttg	3540
atgaagactt	atataatcaat	gccaatgacc	aggcgatcgc	ttatgtgatg	ctcggaaata	3600
aagcccaaac	tgtactaaat	ggcaatttaa	aggtagactt	tatggaaaca	gaaaattcca	3660
gcagtgtgaa	aaaacaaaaa	gcgttagtag	caaaagtgtc	tcagagggaa	gagatggtta	3720
aaaaatgtct	tggagaactt	acagaagtct	gcaaactctc	ggggaaagtt	tttgggtgtc	3780
attacttcaa	tatttttaat	accgtcactc	tcaagaagct	tgcagaatct	ttatcttctg	3840
atcctgaggt	tttgcctcaa	attgatgggt	ttactgaaga	caaactggaa	aaatatgggtg	3900
cggaggtgat	ttcagrtatta	cagaaatact	ctgaatggac	atcgccagct	gaagacagtt	3960
ccccagggat	aagccgtgtcc	agcagcagag	gccccggaag	aagtgccgct	gaggagcttg	4020
acgaggaaat	accggtatct	tcccactact	ttgcaagtaa	aaccagaaat	gaaagggaaga	4080
ggaaaaagat	gccagcctcc	caaaggctct	agaggagaaa	aactgcttcc	agtgggtcca	4140
aggcaagggt	ggggctctgcc	acatgtagaa	agatatcttc	caaaacgaaa	tcctccagca	4200
tcattggatc	cagttcagcc	tcacatactt	ctcaagcgac	atcaggagcc	aatagcaaat	4260
tggggattat	ggctccaccg	aagcctataa	atagaccgtt	tcttaagcct	tcatatgcat	4320
tctcataaca	accgaatctc	aatgtacata	gaccctcttt	cttgtttgtc	agcatctgac	4380
catctgtgac	tataaagctg	ttattcttgt	tataccaaaa	aaaaaaaaaa	aaaaaaa	4437

<210> 36

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000057

<400> 36

taagccttca tatgcattct cataacaacc gaatctcaat gtacatagac cctcttttctt 60

<210> 37
 <211> 2016
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000060

<400> 37
 gccagctgga gcgttttcgg ggctgtaaag ggagaatggc gcatgcgcac attcagggcg 60
 gaaggcgcg taagragcaga tttgtggtct gcattatgtc tggagccaga agtaagcttg 120
 ctcttttctt ctgcggctgt tacgtggttg ccctgggagc ccacaccggg gaggagagcg 180
 tggctgacca tcacgaggct gaatattatg tggctgccgt gtatgagcat ccatccatcc 240
 tgagtctgaa ccctctggct ctcatcagcc gccaaagaggc cttggagctc atgaaccaga 300

accttgacat ctatgaacag caagtgatga ctgcagccca aaaggatgta cagattatag 360
 tgtttccaga agatggcatt catggattca actttacaag aacatccatt tatccatttt 420
 tggacttcat gccgtctccc caggtgggtca ggtggaaccc atgcctggag cctcaccgct 480
 tcaatgacac agaggtgctc cagcgctga gttgtatggc catcagggga gatattgttct 540
 tgggtggccaa tcttgggaca aaggagcctt gtcatagcag tgacccaagg tgcccaaaag 600
 atgggagata ccagttcaac acaaagtctg tgttcagcaa taatggaacc ctgtgtgacc 660
 gctaccgtaa acacaacctc tactttgagg cagcattcga tgttcctctt aaagtggatc 720
 tcatcacctt tgataccccc tttgctggca ggtttggcat ctccacatgc ttgtgatatat 780
 tgttctttga ccctgccatc agagtctca gagactacaa ggtgaagcat gtgtgtgacc 840
 caactgcctg gatgaaccag ctcccactct tggcagcaat tgagattcag aaagcttttg 900
 ctgttgccct tggcatcaac gttctggcag ctaatgtcca ccaccagtt ctggggatga 960
 caggaagtgg catacacacc cctctggagt ccttttggtta ccatgacatg gaaaatccca 1020
 aaagtcaact tataattgcc caggtggcca aaaatccagt gggctctcatt ggtgcagaga 1080
 atgcaacagg tgaaacggac ccattccata gtaagttttt aaaaattttg tcaggcgatc 1140
 cgtactgtga gaaggatgct caggaagtcc actgtgatga ggccaccaag tggaaacgtga 1200
 atgctcctcc cacatttcac tctgagatga tgtatgacaa ttccaccctg gtccctgtct 1260
 ggggaaaggga aggcctatctc cagctctgtt ccaatggcct ctgctgttat ttactttacg 1320
 agaggcccac ctatccaaa gagctgtatg ccctgggggt ctttgatggg ctccacacag 1380
 tacatggcac ttactacatc caagtgtgtg ccctggtcag gtgtgggggt ctgggcttcg 1440
 acacctgcgg acaggaaatc acagagcca cgggatatt tgagtttcac ctgtggggca 1500
 acttcagtac ttcttatatc tttcctttgt ttctgacctc agggatgacc ctagaagtcc 1560
 ctgaccagct tggctgggag aatgaccact atttcttgag gaaaagtagg ctgtcctctg 1620
 ggctgggtgac ggcggctctc tatgggcgct tgtatgagag ggactaggaa aagtgtgtgg 1680
 tctgtggggc ggaactctggc catcatgttg acagccttg acttccacag gctacaagcc 1740
 ctgggaccat ctttctgcct taagggcagg agccacttc tgtggcacca gattccaccc 1800
 tgggaactgt ggaaaaagta ggagaggcag attccctcag tgtcttctc ttaaacctca 1860
 atcatcgaga cattaggggg tattttctgt tcacatttat ctttttcaag ccacatcttc 1920
 ctctaacaaa tctctcagta tgcgattggt ctcaagctaa aacaaaaata aatgtcagtt 1980
 tatattttac acatccaaaa aaaaaaaaaa aaaaaa 2016

<210> 38
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000060

<400> 38
 tcctctaaca aatctctcag tatgcgattg gtctcaagct aaaacaaaaa taaatgtcag 60

<210> 39
 <211> 811
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_000269

<400> 39

```

gcagaagcgt tccgtgcgtg caagtgcgtg gaaccacgtg ggtcccgggc gcgtttcggg 60
tgctggcggc tgcagccgga gttcaaacct aagcagctgg aaggaacctat ggccaactgt 120
gagcgtacct tcattgcgat caaaccagat ggggtccagc ggggtcctgt gggagagatt 180
atcaagcgtt ttgagcagaa aggattccgc cttgttggtc tgaaattcat gcaagcttcc 240
gaagatcttc tcaaggaaaca ctacgttgac ctgaaggacc gtccattctt tgcgggcctg 300
gtgaaataca tgcactcagg gccggtagtt gccatggtct gggaggggct gaatgtggtg 360
aagacgggcc gagtcacgtc cggggagacc aaccctgcag actccaagcc tggggaccatc 420
cgtggagact tctgcataca agttggcagg aacattatac atggcagtga tctctgtggag 480
agtgcagaga aggagatcgg cttgtgggtt caccctgagg aactggtaga ttacacgagc 540
tgtgtcaga actggatcta tgaatgacag gagggcagac cacattgctt ttcacatcca 600
tttccctcc ttcctcatggg cagaggacca ggctgttaga aatctagtta ttacaggaa 660
cttcatcata atttggaggg aagctcttgg agctgtgagt tctccctgta cagtgttacc 720
atccccgacc atctgattaa aatgcttcc cccagcatag gattcattga gttgggttact 780
tcatttgtt gcattgcttt tttttccttc t 811

```

<210> 40

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000269

<400> 40

```

gtctgaaatt catgcaagct tccgaagatc ttctcaagga acactacgtt gacctgaagg 60

```

<210> 41

<211> 2338

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000291

<400> 41

```

agcgcacgtc ggcagtcggc tccctcgttg accgaatcac cgacctctct ccccagctgt 60
atttccaaaa tgcgcgtttc taacaagctg acgctggaca agctggacgt taaagggaag 120
cgggtcggtta tgagagtcga cttcaatgtt cctatgaaga acaaccagat aacaaacaac 180
cagaggatta aggctgctgt cccaagcatc aaattctgct tggacaatgg agccaagtcg 240
gtagtcctta tgagccacct aggccggcct gatggtgtgc ccatgcctga caagtactcc 300
ttagagccag ttgctgtaga actcaaatct ctgctgggca aggatgttct gtctctgaag 360
gactgtgtag gcccagaagt ggagaaagcc tgtgccaaac cagctgctgg gtctgtcatc 420
ctgctggaga acctccgctt tcatgtggag gaagaaggga agggaaaaga tgcttctggg 480
aacaagggtta aagccgagcc agccaaaata gaagctttcc gagcttccact ttccaagcta 540
gggatgtct atgtcaatga tgcttttggc actgctcaca gagcccacag ctccatggta 600
ggagtcaatc tgccacagaa ggctgggtggg tttttgatga agaaggagct gaactacttt 660
gcaaaggcct tggagagccc agagcgaccc ttcttgcca tcctgggcgg agctaaagtt 720
gcagacaaga tcagctcat caataatatg ctggacaaag tcaatgagat gattattggg 780
ggtggaatgg cttttacctt ccttaagggt ctcaacaaca tggagattgg cacttctctg 840
ttgatgaag agggagccaa gattgtcaaa gacctaatgt ccaaagctga gaagaatggt 900
gtgaagatta ccttgccctgt tgactttgtc actgctgaca agtttgatga gaatgccaaag 960
actggccaag ccaactgtggc ttctggcata cctgctggct ggatgggctt ggactgtggt 1020
cctgaaagca gcaagaagta tgctgaggct gtcactcggg ctaagcagat tgtgtggaat 1080
ggtcctgtgg gggatatttga atgggaagct tttgcccggg gaaccaaagc tctcatggat 1140
gaggtggtga aagccacttc taggggctgc atcaccatca taggtggtgg agacactgcc 1200
acttgctgtg ccaaatggaa cacggaggat aaagtcagcc atgtgagcac tgggggtggt 1260
gccagtttgg agctcctgga aggtaaaagtc cttcctgggg tggatgctct cagcaatatt 1320

```

```

tagtactttc ctgcctttta gttcctgtgc acagccccta agtcaactta gcattttctg 1380
catctccact tggcatttagc taaaaccttc catgtcaaga ttcagctagt ggccaagaga 1440
tgcagtgcc a ggaaccctta aacagttgca cagcatctca gctcatcttc actgcaccct 1500
ggatttgcac acatttcttca agatcccat tgaatttttt agtgactaaa ccattgtgca 1560
ttctagagtg catatatatta ttttttgcct gttaaaaaga aagtgagcag tgttagctta 1620
gttctctttt gatgtagggtt attatgatta gctttgtcac tgtttcacta ctcagcatgg 1680
aaacaagatg aaatttccatt tgtaggtagt gagacaaaat tgatgatcca ttaagtaaac 1740
aataaaagtg tccattgaaa ccgtgatttt ttttttttct ctgtcatact ttgttaggaa 1800
gggtgagaat agaattcttga ggaacggatc agatgtctat attgctgaat gcaagaagtg 1860
gggcagcagc agtgagagaga tgggacaatt agataaatgt ccatttctta tcaagggcct 1920

```

```

actttatggc agacattgtg ctagtgcttt tattctaact tttattttta tcagttacac 1980
atgatcataa tttaaaaagt caaggcttat aacaaaaaag cccagccca ttcctcccat 2040
tcaagattcc cactccccag aggtgaccac tttcaactct tgagtttttc aggtatatac 2100
ctccatgttt ctaagtaata tgcttatatt gttcacttcc ttttttttta ttttttaaag 2160
aaatctatct cataccatgg aggaaggctc tggtccacat atatttccac ttcttcatc 2220
tctcggtata gttttgtcac aattatagat tagatcaaaa gtctacataa ctaatacagc 2280
tgagctatgt agtatgctat gattaaattt acttatgtaa aaaaaaaaaa aaaaaaaaaa 2338

```

<210> 42

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000291

<400> 42

```

acttagcatt ttc tgcacat ccacttggca ttagctaaaa ccttccatgt caagattcag 60

```

<210> 43

<211> 787

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000363

<400> 43

```

ctgaagggtca cccggggcggc cccctcactg accctccaaa cgccctgtgc ctcgcctgc 60
ctcctgccat tcccggcctg agtctcagca tggcggatgg gagcagcgat gcggctaggg 120
aacctcgccc tgcaccagcc ccaatcagac gccgctcctc caactaccgc gcttatgcca 180
cggagccgca cgc caagaaa aaatctaaga tctccgcctc gagaaaattg cagctgaaga 240
ctctgtgtgt gcaagattgca aagcaagagc tggagcgaga ggcggaggag cggcgcgagg 300
agaaggggag cgc tctgagc acccgtgccc agccgtgga gttgaccggg ctgggcttcg 360
cggagctgca gga.cttgtgc cgacagctcc acgcccgtgt ggacaagggt gatgaagaga 420
gatacgacat agaggcaaaa gtcaccaaga acatcacgga gattgcagat ctgactcaga 480
agatctttga ccttcgagggc aagtttaagc ggcccaccct gcggagagtg aggatctctg 540
cagatgccat gatgcaggcg ctgctggggg cccgggctaa ggagtccctg gacctgcggg 600
cccacctcaa gcaagggtgaag aaggaggaca ccgagaagga aaaccgggag gtgggagact 660
ggcggaagaa catcgatgca ctgagtggaa tggagggccg caagaaaaag tttgagagct 720
gagccttctc gcctactgcc cctgccctga ggagggccac tgaggaataa agcttctctc 780
tgagctg 787

```

<210> 44

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000363

<400> 44
tgtggaCaag gtggatgaag agagatacga catagaggca aaagtcacca agaacatcac 60

<210> 45
<211> 1263
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 45
ggcacgagac cttcagcgcc tcggctccag cgccatggcg ccctccagga agttcttcgt 60
tgggggaaac tgggaagatga acggggcgga gcagagtctg ggggagctca tcggcactct 120
gaacgcggcc aaggtgccgg ccgacaccga ggtgggttgt gctcccccta ctgcctatat 180
cgacttcgcc cggcagaagc tagatcccaa gattgctgtg gctgcgcaga actgctacaa 240
agtgcataat ggggctttta ctggggagat cagccctggc atgatcaaag actgcggagc 300
cacgtgggtg gtccctggggc actcagagag aaggcatgtc tttggggagt cagatgagct 360
gattgggcag aaagtggccc atgctctggc agagggactc ggagtaatcg cctgcattgg 420
ggagaaacta gatgaaaggg aagctggcat cactgagaag gttgttttcg agcagacaaa 480
ggatcatcgca gataacgtga aggactggag caaggctcgc ctggcctatg agcctgtgtg 540
ggccat tggc actggcaaga ctgcaacacc ccaacaggcc caggaagtac acgagaagct 600
ccgaggatgg ctgaagtcca acgtctctga tgcgggtggc cagagcaccg gtatcattta 660
tggaggctct gtgactgggg caacctgcaa ggagctggcc agccagcctg atgtggatgg 720
cttcct tgtg ggtggtgctt ccctcaagcc cgaattcgtg gacatcatca atgccaaaaca 780
atgagcccca tccatcttcc ctacccttcc tgccaagcca gggactaagc agcccagaag 840
cccagt aact gccctttccc tgcataatgct tctgatgggtg tcatctgctc cttcctgtgg 900
cctcat ccaa actgtatctt cctttactgt ttatatcttc accctgtaat ggttggggacc 960
aggccaatcc cttctccact tactataatg gttggaacta aacgtcacca aggtggcctc 1020
tccttggtcg agagatggaa ggcgtgggtg gatttgctcc tgggttcctt aggccctagt 1080
gagggcagaa gagaaaccat cctctccctt cttacaccgt gagggcaaga tcccctcaga 1140
aggcaggagt gctgccctct cccatggtgc ccgtgcctct gtgctgtgta tgtgaaccac 1200
ccatgtgagg gaataaacct ggcactagga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
aaa 1263

<210> 46
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000365

<400> 46
tatcttcacc ctgtaatggc tgggaccagg ccaatccctt ctccacttac tataatgggtt 60

<210> 47
<211> 1616
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000582

<400> 47
ctccctgtgt tgggtggagga tgtctgcagc agcatttaaa ttctgggagg gcttgggtgt 60
cagcagcagc aggaggaggc agagcacagc atcgctggga ccagactcgt ctcaggccag 120
ttgcagcctt ctcagccaaa cgccgaccaa ggaaaactca ctaccatgag aattgcagtg 180
atttgctttt gcctcctagg catcacctgt gccataccag ttaaacaggc tgattctgga 240
agttctgagg aaaagcagct ttacaacaaa taccagatg ctgtggccac atggctaaac 300


```

cctgacccat ctcagaagca gaatctccta gccccacaga cccttccaag taagtccaaC 360
gaaagccatg accacatgga tgatatggat gatgaagatg atgatgacca tgtggacagC 420
caggactcca ttgactcgaa cgactctgat gatgtagatg aactgatga ttctcaccag 480
tctgatgagt ctcaccattc tgatgaatct gatgaactgg tcaactgattt tcccacgggaC 540
ctgccagcaa ccgaagtttt cactccagtt gtccccacag tagacacata tgatggccga 600
ggtgatagtg tggtttatgg actgaggtca aaatctaaga agtttcgcag acctgacatC 660
cagtaccctg atgctacaga cgaggacatc acctcacaca tggaaagcga ggagttgaatC 720
ggtgcataca aggccatccc cgttgccag gacctgaacg cgccttctga ttgggacagC 780
cgtgggaagg acagttatga aacgagtcag ctggatgacc agagtgtga aacccacagC 840
cacaagcagt ccagattata taagcggaaa gccaatgatg agagcaatga gcattccgatC 900
gtgattgata gtcaggaact ttccaaagtc agccgtgaat tccacagcca tgaatttcaC 960
agccatgaag atatgctggt tgtagacccc aaaagtaagg aagaagataa acacctgaaa 1020
tttctgattt ctcattgaatt agatagtgc tcttctgagg tcaattaaaa ggagaaaaaa 1080
tacaatttct cactttgcat ttagtcaaaa gaaaaaatgc tttatagcaa aatgaaaga 1140
aacatgaaat gcttctttct cagttttattg gttgaatgtg tatctatttg agtctggaaa 1200
taactaatgt gtttgataat tagtttagtt tgtggcttca tggaaactcc ctgtaaaacta 1260
aaagcttcag ggttatgtct atgttcattc tatagaagaa atgcaaaacta tcaactgtatC 1320
ttaatatattg ttattctctc atgaatagaa atttatgtag aagcaaacia aatactttta 1380
cccacttaaa aagagaatat aacattttat gtcactataa tcttttgttt tttaagttag 1440
tgtatatattt gttgtgatta tctttttgtg gtgtgaataa atcttttatc ttgaatgtaa 1500
taagaatttg gtggtgtcaa ttgcttattt gttttccac ggttgtccag caattaataa 1560
aacataacct tttttactgc ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1616

```

<210> 48
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_0 00582

```

<400> 48
ggtggtgtca attgcttatt tgttttccca cggttgtcca gcaattaata aaacataacC 60

```

<210> 49
 <211> 1666
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_0 00584

```

<400> 49
ctccataagg cacaaacttt cagagacagc agagcacaca agcttctagg acaagagcc a 60
ggaagaaacc accggaagga accatctcac tgtgtgtaaa catgacttcc aagctggcc g 120
tggctctctt ggcagccttc ctgatttctg cagctctgtg tgaagggtgca gttttgcca a 180
ggagtgtctaa agaacttaga tgtcagtgc taaagacata ctccaaacct ttccacccc a 240
aatttatcaa agaactgaga gtgattgaga gtggaccaca ctgcgccaac acagaaatt a 300
ttgtaaagct ttctgatgga agagagctct gtctggaccc caaggaaaac tgggtgcaga 360
gggttgtgga gaagtttttg aagagggctg agaattcata aaaaaattca ttctctgtg g 420
tatccaagaa tcagtgaaga tgccagtga acttcaagca aatctacttc aacacttca t 480
gtattgtgtg ggtctgttgt agggttgcca gatgcaatac aagattcctg gttaaattt g 540
aatttcagta aacaatgaat agtttttcat tgtaccatga aatatccaga acatactta t 600
atgtaaagta ttatttattt gaatctacaa aaaacaacia ataattttta aatataagga 660
ttttcctaga tattgcacgg gagaatatac aaatagcaaa attgaggcca agggccaaga 720
gaatatccga actttaattt caggaattga atgggtttgc tagaatgtga tatttgaag c 780
atcacataaa aatgatggga caataaattt tgccataaag tcaaatttag ctggaaatc c 840
tggatttttt tctgttaaat ctggcaaccc tagtctgcta gccaggatcc acaagtcct t 900
gttccactgt gccttggttt ctcttttatt tctaagtggg aaaagtatta gccaccatc t 960
tacctcacag tgatgttgtg aggacatgtg gaagcacttt aagttttttc atcataaca t 1020
aaattatttt caagtgtaac ttattaacct atttattatt tatgtattta tttaagcat c 1080

```

```

aaatatttgt gcaagaattt ggaaaaatag aagatgaatc attgat tgaa tagttataaa 1140
gatgttatag taaattttatt ttatttttaga tattaaatga tgttttatta gataaatttc 1200
aatcagggtt ttttagattaa acaaacaaac aattgggtac ccagttaaat tttcatttca 1260
gataaacaac aaataattttt ttagtataag tacattattg tttatc tgaa attttaattg 1320
aactaacaat cctagtttga tactcccagt cttgtcattg ccagct gtgt tggtagtgct 1380
gtgttgaat t acggaataat gagttagaac tattaaaaca gccaaa actc cacagtcaat 1440
attagtaat t tcttgctggg tgaacttgt ttattatgta caaata gatt cttataatat 1500
tatttaaatg actgcatttt taaatacaag gctttatatt tttaac tttt agatgttttt 1560
atgtgctct c caaatttttt ttactgtttc tgattgtatg gaaata taaa agtaaatatg 1620
aaacatttta a aatataattt gttgtcaaag taaaaaaaa aaaaaa 1666

```

<210> 50

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000584

<400> 50

```

tggtagtgc t gtgttgaatt acggaataat gagttagaac tattaaaaca gccaaaactc 60

```

<210> 51

<211> 172 2

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000599

<400> 51

```

ggggaaaaga gctaggaaag agctgcaaag cagtgtgggc tttttccctt tttttgctcc 60
ttttcattac cctcctccg ttttcacct tctccggact tcgcgt agaa cctgcgaatt 120
tcgaagagga ggtggcaaag tgggagaaaa gaggtgttag ggtttggggt ttttttgttt 180
ttgtttttgt tttttaattt cttgatttca acattttctc ccacct tctc ggctgcagcc 240
aacgcctct t acctgttctg cggcgccgcg caccgctggc agctga gggg tagaaagcgg 300
ggtgtatttt agattttaag caaaaatttt aaagataaat ccattt ttct ctcccacccc 360
caacgccat c tccactgcat ccgatctcat tatttcgggtg gttgct tggg ggtgaacaat 420
tttgtggct t tttttccctt ataattctga cccgctcagg cttgaggggt tctccggcct 480
ccgctcactg cgtgcacctg gcgctgccc gcttccccca acctgt tgca aggctttaat 540
tcttgcaact gggacctgct cgcaggcacc ccagccctcc acctct ctct acatttttgc 600
aagtgtctgg gggagggcac ctgctctacc tgccagaaat tttaaa acaa aaacaaaaac 660
aaaaaaatct cccggggccc tcttgcccc tttatccctg cactct cgct ctccctgccc 720
accccgagggt aaaggggggc actaagagaa gatggtgttg ctcaaccgcg tctcctctgct 780
gctggccgc c tatgcggggc cggcccagag cctgggctcc ttcgtg cact gcgagccctg 840
cgacgagaaa gccctctcca tgtgcccccc cagccccctg ggctgc gagg tggtaagga 900
gccgggctgc ggctgctgca tgacctgcgc cctggccgag gggcagtcgt gccggcgtcta 960
caccgagcgc tgcgcccagg ggctgcgctg cctcccccg caggac gagg agaagccgct 1020
gcacgccctg ctgcacggcc gcgggggttg cctcaacgaa aagagc tacc gcgagcaagt 1080
caagatcgag agagactccc gtgagcacga ggagcccacc acctct gaga tggccgagga 1140
gacctactcc cccaagatct tccggcccaa acacaccgac atctcc gagg tgaaggctga 1200
agcagtgaag aaggaccgca gaaagaagct gaccagtcct aagttt gtcg ggggagccga 1260
gaacactgcc ccccccgga tcatctctgc acctgagatg agacaggagt ctgagcaggg 1320
ccctgcccgc agacacatgg aggcctccct gcaggagctc aaagc cagcc cacgcatggt 1380
gcccctgtgc gtgtacctgc ccaattgtga ccgcaaagga ttcta caaga gaaagcagtg 1440
caaaccttcc cgtggccgca agcgtggcat ctgctgggtg gtgga caagt acgggatgaa 1500
gctgccaggc atggagtacg ttgaaggggg ctttcagtgc cacac cttcg acagcagcaa 1560
cgttgagtga tgcgtcccc cccaacctt cctcaccct ctccca cccc cagccccgac 1620
tccagccagc gcctccctcc accccaggac gccactcatt tcactc catt taagggaaaa 1680
atatatatct atctatttga ggaaaaaaa aaaaaaaaa aa 1722

```

<210> 52
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000599

<400> 52
 ccaggaCgcc actcatttca tctcatttaa gggaaaaata tataTctatc tatttgagga 60

<210> 53
 <211> 704
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000735

<400> 53
 gcagttactg agaactcata agacgaagct aaaatccctc ttcggatcca cagtcaaccg 60
 ccctgaacac atcctgcaaa aagcccagag aaaggagcgc catggattac tacagaaaat 120
 atgcagctat ctttctggtc acattgtcgg tgtttctgca tgttctccat tccgctcctg 180
 atgtgcagga ttgccagaa tgcacgctac aggaaaacc attcttctcc cagccgggtg 240
 ccccaatctc tcagtgcacg ggctgctgct tctctagagc atatccact ccactaaggc 300
 ccaagaagac gatgttggtc caaaagaacg tcacctcaga gtccacttgc tgtgtagcta 360
 aatcatataa cagggtcaca gtaatggggg gtttcaaagt ggagaaccac acggcgtgcc 420
 actgcagtac ttgttattat cacaaatctt aaatgtttta ccaagtgtctg tcttgatgac 480
 tgctgatttt ctggaatgga aaattaagtt gtttagtggt tatggctttg tgagataaaa 540
 ctctcctttt ccttaccata ccactttgac acgcttcaag gataTactgc agctttactg 600
 ccttccTcct tatcctacag tacaatcagc agtctagttc ttttCatttg gaatgaatac 660
 agcattaagc ttgttccact gcaaataaag ccttttaaat catc 704

<210> 54
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000735

<400> 54
 tgagataaaa ctctcctttt ccttaccata ccactttgac acgcttcaag gatatactgc 60

<210> 55
 <211> 1342
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000799

<400> 55
 cccggagccg gaccggggcc accgcgcccg ctctgctccg acacCgcgcc ccctggacag 60
 ccgccctctc ctccaggccc gtggggctgg ccctgcaccg ccgagcttcc cgggatgagg 120
 gccccgggtg tggtcacccg gcgcgccccg ggtcgtgag ggacCccggc caggcgcgga 180
 gatgggggtg cacgaatgtc ctgcctggct gtggcttctc ctgtCcttgc tgtcgtctcc 240
 tctgggcctc ccagtcctgg gcgccccacc acgcctcatc tgtgacagcc gagtcttgga 300
 gaggtacctc ttggaggcca aggaggccga gaatatcag acgggctgtg ctgaacactg 360
 cagcttgaat gagaatatca ctgtcccaga caccaaagtt aatttctatg cctggaagag 420

```

gatggagggtc gggcagcagg ccgtagaagt ctggcagggc ctggccctgc tgtcgggaagc 480
tgtccctgcgg ggccaggccc tgttgggtcaa ctcttcccag ccgtgggagc ccctgcagct 540
gcatgtggat aaagccgtca gtggccttcg cagcctcacc actctgcttc gggctctgcg 600
agcccagaag gaagccatct cccctccaga tgcggcctca gctgctccac tccgaacaat 660
cactgctgac actttccgca aactcttccg agtctactcc aatttcctcc ggggaaagct 720
gaagctgtac acaggggagg cctgcaggac aggggacaga tgaccagggtg tgtccacctg 780
ggcatatcca ccacctccct caccaacatt gcttgtgcc aacctcccc cgccactcct 840
gaaccccgtc gaggggctct cagctcagcg ccagcctgtc ccatggacac tccagtgcc 900
gcaatgacat ctcaggggccc agaggaactg tccagagagc aactctgaga tctaaggatg 960
tcacagggccc aacttgaggg cccagagcag gaagcattca gagagcagct ttaaactcag 1020
ggacagagcc atgctgggaa gacgcctgag ctactcggc accctgcaaa atttgatgcc 1080
aggacacgct ttggaggcga ttacctgtt ttcgcacct ccatcaggga caggatgacc 1140
tggagaactt aggtggcaag ctgtgacttc tccaggtctc acgggcatgg gcactccctt 1200
ggtggcaaga gcccccttga caccggggtg gtgggaacca tgaagacagg atgggggctg 1260
gcctctggct ctcatggggt ccaagttttg tgtattcttc aacctcattg acaagaactg 1320
aaaccaccaa aaaaaaaaaa aa 1342

```

<210> 56

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000799

<400> 56

```

tcattgggggtc caagttttgt gtattcttca acctcattga caaggaactga aaccaccaa 60

```

<210> 57

<211> 2722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000917

<400> 57

```

gagcggggctg agggtaggaa gtagccgctc cgagtggagg cgactggggg ctgaagagcg 60
cgccggccctc tcgtcccact ttccagggtg gtgatcctgt aaaattaaat cttccaagat 120
gatctgggtat atattaatta taggaattct gcttcccag tctttggctc atccaggctt 180
ttttacttca attggtcaga tgactgattt gatccatact gagaaagatc tgggtgacttc 240
tctgaaagat tatattaagg cagaagagga caagttagaa caaataaaaa aatgggcaga 300
gaagttagat cggctaacta gtacagcgac aaaagatcca gaaggatttg ttgggcatcc 360
agtaaattgca ttcaaattaa tgaaacgtct gaatactgag tggagtgagt tggagaatct 420
ggtccttaag gatattgtcag atggctttat ctctaacct accattcaga gaccagtact 480
ttctaatgat gaagatcagg ttggggcagc caaagctctg ttacgtctcc aggataccta 540
caatttggat acagatacca tctcaaaggg taatcttcca ggaagtgaac acaaattctt 600
tctaacggct gaggactgct ttgagttggg caaagtggcc tatacagaag cagattatta 660
ccatacggaa ctgtggatgg aacaagccct aaggcaactg gatgaaggcg agatttctac 720
catagataaa gtctctgttc tagattatct gagctatgcg gtaatatcagc agggagacct 780
ggataaggca cttttgtctc caaagaagct tcttgaacta gatcctgaac atcagagagc 840
taatggtaac ttaaaatatt ttgagtatat aatggctaaa gaaaagatg tcaataagtc 900
tgcttcagat gaccaatctg atcagaaaac tacaccaaag aaaaagggg ttgctgtgga 960
ttactgcca gagagacaga agtacgaaat gctgtgccgt ggggagggta tcaaaatgac 1020
ccctcggaga cagaaaaaac tcttttgccg ctaccatgat ggaaaccgta atcctaaatt 1080
tattctggct ccagctaaac aggaggatga atgggacaag cctcgtatta ttcgcttcca 1140
tgatattatt tctgatgcag aaattgaaat cgtcaaagac ctacgcaaac caaggctgag 1200
ccgagctaca gtacatgacc ctgagactgg aaaattgacc acagcacagt acagagtatc 1260
taagagtgcc tggctctctg gctatgaaaa tcctgtgggt tctcgaatta atatgagaat 1320
acaa gatcta acaggactag atgtttccac agcagaggaa ttaacaggtag caaattatgg 1380
agttggagga cagtatgaac cccattttga ctttgcacgg aaagatgagc cagatgcttt 1440

```

caaa	gagctg	gggacaggaa	atagaattgc	tacatggctg	ttttatatga	gtgatgtgtc	1500
tgcaggagga	gccactgttt	ttcctgaagt	tggagctagt	gtttggccca	aaaaaggaac		1560
tgctgttttc	tggtataatc	tgtttgccag	tggagaagga	gattatagta	cacggcatgc		1620
agcctgtcca	gtgctagtgt	gcaacaaatg	ggtatccaat	aaatggctcc	atgaacgtgg		1680
acaagaattt	cgaagacctt	gtacgttgtc	agaattggaa	tgacaaacag	gcttcccttt		1740
ttctcctatt	gttgtaacct	tatgtgtctg	atatacacat	ttccatagtc	ttaaactttca		1800
ggagttttaca	attgactaac	actccatgat	tgattcagtc	atgaacctca	tcccatgttt		1860
catctgtgga	caattgctta	ctttgtgggt	tcttttaaaa	gtaacacgaa	atcatcatat		1920
tgcataaaaac	cttaaagtgc	tggtgggtatc	acagaagaca	aggcagagtt	taaagtgagg		1980
aatttttata	ttaaagaact	ttttgggttg	ataaaaacat	aatttgagca	tccagtttta		2040
gtattttcact	acatctcagt	tggtgggtgt	taagctagaa	tgggctgtgt	gataggaaac		2100
aaatgcctta	cagatgtgcc	taggtgttct	gtttacctag	gtcttactc	tgttttctgg		2160

atctgaagac	tagtaataaa	ctaggacact	aactgggttc	catgtgattg	ccctttcata	2220
tgatcttcta	agttgatttt	tttctctcca	agtctttttt	aaagaaagta	tactgtattt	2280
taccaacccc	ctctcttttc	ttttagctcc	tctgtgggtga	attaaacgta	cttgagttta	2340
aataatttcga	tttttttttt	ttttttaatg	gaaagtcttg	cataacaaca	ctgggccttc	2400
ttaactaaaa	tgctcaccac	ttagcctggt	tttttatccc	ttttttaaaa	tgacagatga	2460
ttttgttcag	gaattttgct	gtttttctta	gtgctaatac	cttgccctct	attcctgcta	2520
cagcagggtg	gtaatatattg	cattctgatt	aaatactgtg	cttagggaga	ctggaagttt	2580
aaaaatgtac	aagtcctttc	agtgatgagg	gaattgattt	tttttaaaag	tctttttctt	2640
agaaagccaa	aatgtttgtt	tttttaagat	tctgaaatgt	gttggtgacaa	caatgacctt	2700
tttatgatct	taaatctttt	tt	2722			

<210> 58

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000917

<400> 58

tcttactctg	ttttctggat	ctgaagacta	gtaataaact	aggacactaa	ctgggttcca	60
------------	------------	------------	------------	------------	------------	----

<210> 59

<211> 3236

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001109

<400> 59

gacccggcca	tgcgcggcct	cgggctctgg	ctgctgggcg	cgatgatgct	gcctgcgatt	60
gccccagcc	ggccctgggc	cctcatggag	cagtatgagg	t<gtgttgcc	gcggcgtctg	120
ccaggccccc	gagtcgcgcg	agctctgccc	tcccacttgg	g<ctgcaccc	agagaggggtg	180
agctacgtcc	ttggggccac	agggcacaa	ttcacccctc	a<ctgcggaa	gaacaggggac	240
ctgctgggtt	ccggctacac	agagacctat	acggctgcca	a<tggtccga	ggtgacggag	300
cagcctcgcg	ggcaggacca	ctgcttatac	cagggccacg	tagaggggta	cccggactca	360
gccgcagcc	tcagcacctg	tgccggcctc	aggggtttct	t<caggtggg	gtcagacctg	420
cacctgatcg	agcccttgga	tgaagggtgc	gagggcggac	ggcacgccgt	gtaccaggct	480
gagcacctgc	tcagacgggc	cgggacctgc	ggggtcagcg	a<cgacagcct	gggcagcctc	540
ctgggacccc	ggacggcagc	cgtcttcagg	cctcggcccg	gggactctct	gccatcccga	600
gagacccgct	acgtggagct	gtatgtgggtc	gtggacaatg	cagagttcca	gatgctgggg	660
agcgaagcag	ccgtgcgtca	tcgggtgctg	gaggtgggtga	a<tcacgtgga	caagctatat	720
cagaaactca	acttcctgtg	ggtcctgggtg	ggcctggaga	t<ttggaatag	tcaggacagg	780
ttccacgtca	gccccgaccc	cagtgtcaca	ctggagaacc	t<cctgacctg	gcaggcacgg	840
caa<ggacac	ggcggcacct	gcatgacaac	gtacagctca	t<cacgggtgt	cgacttcacc	900
gggactactg	tgggggtttgc	caggggtgtcc	gccatgtgct	c<ccacagctc	aggggctgtg	960
aac<aggacc	acagcaagaa	ccccgtgggc	gtggcctgca	c<catggcca	tgagatgggc	1020

```

cacaacctgg gcatggacca tgatgagaac gtccagggct gCcgctgcca ggaacgcttc 1080
gaggccggcc gctgcatcat ggcaggcagc attggctcca gtttccccag gatgttcagt 1140
gactgcagcc aggcctacct ggagagcttt ttggagcggc cgcagtcggg gtgcctcgcc 1200
aacgcccctg acctcagcca cctggtgggc ggccccgtgt gtgggaacct gtttgtggag 1260
cgtggggagc agtgcgactg cggccccccc gaggactgcc ggaaccgctg ctgcaactct 1320
acca.cctgcc agctggctga gggggcccag tgtgcgcacg gtacctgctg ccaggagtgc 1380
aaggTgaagc cggctggtga gctgtgccgt cccaagaagg acatgtgtga cctcgaggag 1440
ttctgtgacg gccggcacc tgagtgcccg gaagacgcct tccaggagaa cggcacgccc 1500
tgctccgggg gctactgcta caacggggcc tgtcccacac tggcccagca gtgccaggcc 1560
ttctgggggc caggtgggca ggctgccgag gagtccctgt tctcctatga catcctacca 1620
ggctgcaagg ccagccggta cagggctgac atgtgtggcg tctgtcagtg caagggtggg 1680
cagcagcccc tggggcgtgc catctgcatc gtggatgtgt gccacgcgct caccacagag 1740
gatggcactg cgtatgaacc agtgcccag ggcacccggg gtggaccaga gaaggtttgc 1800
tgga.aaggac gttgccagga cttacacgtt tacagatcca gcaactgctc tgcccagtgc 1860
caca.accatg gggTgtgcaa ccacaagcag gagtgccact gccacgcggg ctgggccccg 1920
ccccactgcg cgaagctgct gactgaggtg cagcagcgt cggggagcct cccgtcctc 1980
gtggTggTgg ttctgggtgt cctggcagtt gtgctggTca cctggcagg catcatcgtc 2040
taccgcaaag cccggagcgg catcctgagc aggaacgtgg cTcccaagac cacaatgggg 2100
cgctccaacc ccctgttcca ccaggctgcc agccgcgtgc cggccaaggg cggggctcca 2160
gccc.catcca ggggccccca agagctggtc ccaccaccc aCccgggcca gccgcgccga 2220
cacc.cggcct cctcggtggc tctgaagagg ccgccccctg cTcctccggg cactgtgtcc 2280
agcc.caccct tcccagttcc tgtctacacc cggcaggcac caaagcaggt catcaagcca 2340
acgt.tcgcac ccccagtgcc ccagTcaaa cccggggctg gtgcgggcaa cctggtcca 2400
gctgagggtg ctgttggccc aaaggttgc ctgaagcccc ccatccagag gaagcaagga 2460
gccggagctc ccacagcacc ctaggggggc acctgcgct gtgtggaaat ttggagaagt 2520
tgcggcagag aagccatgcg ttccagcctt ccacggTcca gctagtgcg ctcagcccta 2580
gacc.ctgact ttgcaggctc agctgctgtt ctaacctcag taatgcatct acctgagagg 2640
ctcc.tgctgt ccacgcctc agccaattcc ttctccccgc cttggccacg tgtagcccca 2700
gctgtctgca ggcaccaggc tgggatgagc tgtgtgcttg cgggtgcgtg tgtgtgtacg 2760
tgTc.tccagg tggccgctgg tctcccgtg tgttcaggag gccacatata cagccccctc 2820
cagc.cacacc tgcccctgct ctggggcctg ctgagccggc tgccttgggc acccggttcc 2880
aggcagcaca gacgtggggc atccccagaa agactccatc cCaggaccag gttccccctc 2940
gtgc.tcttcg agagggTgtc agtgagcaga ctcacccca agctcccga tccaggTccc 3000
ctga.tcttgg gctgtttcc catgggattc aagagggaca gccccagctt tgtgtgtgtt 3060
taag.cttagg aatgcccttt atggaaagg ctatgtggga gagtcagcta tcttgtctgg 3120
tttt.cttgag acctcagatg tgtgttcagc agggctgaaa gctttttattc ttttaataatg 3180
agaa.atgtat attttactaa taaattattg accgagttct gtagattctt gttaga 3236

```

<210> 60
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001109

<400> 60
 ctttatggaa agggctatgt gggagagtca gctatcttgt ctgggttttct tgagacctca 60

<210> 61
 <211> 1449
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001124

<400> 61
 ctggratagaa cagctcaagc cttgccactt cgggcttctc actgcagctg ggcttggact 60
 tcggagtttt gccattgccg gtgggacgtc tgagactttc tCcttcaagt acttggcaga 120
 tcactctctt agcagggtct gcgcttcgca gccgggatga agctgggttc cgtcgccctg 180

atgtacctgg	gttcgctcgc	cttcctaggg	gctgacaccg	ctcgggttga	tgtcgcgtcg	240
gagtttccgaa	agaagtggaa	taagtgggct	ctgagtcgtg	ggaagagggg	actgcggatg	300
tccagcagct	accccaccgg	gctcgcgtgac	gtgaaggccg	ggcctgccca	gacccttatt	360
cggccccagg	acatgaagg	tgcctctcga	agccccgaag	acagcagtc	ggatgccgcc	420
cgcacccgag	tcaagcgcta	ccgccagagc	atgaacaact	tccagggcct	ccggagcttt	480
ggctgccgct	tcgggacgtg	cacgggtgcag	aagctggcac	accagatcta	ccagttcaca	540
gataaggaca	aggacaacgt	cgccccagg	agcaagatca	gccccagg	ctacggccgc	600
cggcgccggc	gctccctgcc	cgaggccggc	ccgggtcgga	ctctggtgtc	ttctaagcca	660
caagcacacg	gggctccagc	ccccccgagt	ggaagtgtc	cccactttct	ttaggattta	720
ggcgccccatg	gtacaaggaa	tagtcgcgca	agcatccgc	tgggtgcctcc	cgggacgaag	780
gactttcccga	gcgggtgtgg	gaccgggctc	tgacagccct	gcgagagacc	tgagtccggg	840
agggaccgctc	cggcgccgag	ctctggcttt	gcaagggccc	ctccttctgg	gggttcgct	900
tccttagcct	tgctcaggtg	caagtgtccc	agggggcg	gtgcagaaga	atccgagtgt	960
ttgccaggct	taaggagagg	agaaactgag	aatgaatgc	tgagaccccc	ggagcagggg	1020
tctgagccac	agccgtgtc	gcccacaaac	tgatttctca	cggcgtgtca	ccccaccagg	1080
gcgcaagcct	caactattact	tgaactttcc	aaaacctaaa	gaggaaaagt	gcaatgcgtg	1140
ttgtacatac	agaggtaact	atcaatattt	aagtttgttg	ctgtcaagat	tttttttgta	1200
actttcaaata	tagagatatt	tttgtacgtt	atatattgta	ttaagggcat	tttaaaagca	1260
attatattgt	cctccctat	tttaagacgt	gaatgtctca	gcgaggtgta	aagttgttcg	1320
ccgcgtggaa	tgtgagtgtg	tttgtgtgca	tgaaagagaa	agactgatta	cctcctgtgt	1380
ggaagaagga	aacaccgagt	ctctgtataa	tctatttaca	taaaatgggt	gatatgcgaa	1440
cagcaaacc	1449					

<210> 62

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 62

gaaggaaaca ccgagtctct gtataatcta ttacataaa atgggtgata tgcgaacagc 60

<210> 63

<211> 1619

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001168

<400> 63

ccgccagatt	tgaatcgccg	gaccggttgg	cagaggtggc	ggcgccggca	tgggtgcccc	60
gacgttgccc	cctgcctggc	agccctttct	caaggaccac	cgcactctca	cattcaagaa	120
ctggcccttc	ttggagggtc	gcgcctgcac	cccggagcgg	atggccgagg	ctggcttcat	180
ccactgcccc	actgagaacg	agccagactt	ggcccagtg	ttcttctgct	tcaaggagct	240
ggaaggctgg	gagccagatg	acgaccccat	agaggaacat	aaaaagcatt	cgtccgggtg	300
cgcctttcctt	tctgtcaaga	agcagtttga	agaattaacc	cttgggtgaat	ttttgaaact	360
ggacagagaa	agagccaaga	acaaaattgc	aaaggaaacc	aacaataaga	agaaagaatt	420
tgaggaaact	gcgaagaaag	tgcgccgtgc	catcgagcag	ctggctgcca	tggattgagg	480
cctctggccg	gagctgcctg	gtcccagagt	ggctgcacca	cttccagggt	ttattccctg	540
gtgccaccag	ccttccctgtg	ggccccttag	caatgtctta	ggaaaggaga	tcaacatttt	600
caaattagat	gtttcaactg	tgctcctgtt	ttgtcttgaa	agtggcacca	gaggtgcttc	660
tgcctgtgca	gcgggtgctg	ctggtaacag	tggctgcttc	tctctctctc	tctctttttt	720
gggggtcat	ttttgctgtt	ttgattcccg	ggcttaccag	gtgagaagtg	agggaggaag	780
aaggcagtg	cccttttgct	agagctgaca	gctttgttcg	cgtgggcaga	gccttcacac	840
gtgaatgtgt	ctggacctca	tggtgttgag	gctgtcacag	tcctgagtgt	ggacttgcca	900
ggtgcctgtt	gaatctgagc	tgcaggttcc	ttatctgtca	cacctgtgcc	tcctcagagg	960
acagtttttt	tgttgtgtg	tttttttgtt	tttttttttt	ggtagatgca	tgacttgtgt	1020
gtgatgagag	aatggagaca	gagtcctctg	ctcctctact	gtttaacaac	atggctttct	1080

tat tttgttt	gaattgttaa	ttcacagaat	agcacaaact	acaattaaaa	ctaagcacia	1140
agccattcta	agtcattggg	gaaacgggg	gaacttcagg	tggatgagga	gacagaatag	1200
agt gatagga	agcgtctggc	agatactcct	tttgccactg	ctgtgtgatt	agacaggccc	1260
agt gagccgc	ggggcacatg	ctggccgctc	ctccctcaga	aaaaggcagt	ggcctaaatc	1320
cttttttaaat	gacttggctc	gatgctgtgg	gggactggct	gggctgctgc	aggccgtgtg	1380
tctgtcagcc	caaccttcac	atctgtcacg	ttctccacac	gggggagaga	cgagtcgcgc	1440
ccaggtcccc	gctttctttg	gaggcagcag	ctcccgcagg	gctgaagtct	ggcgtaagat	1500
gatggatttg	attcgccctc	ctccctgtca	tagagctgca	gggtggattg	ttacagcttc	1560
gctggaaacc	tctggaggtc	atctcggtcg	ttctgagaa	ataaaaagcc	tgtcatttc	1619

<210> 64

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001168

<400> 64

ttcacagaat	agcacaaact	acaattaaaa	ctaagcacia	agccattcta	agtcattggg	60
------------	------------	------------	------------	------------	------------	----

<210> 65

<211> 1552

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001216

<400> 65

gcccgtagac	accgtgtgct	gggacacccc	acagtcagcc	gcatggctcc	cctgtgcccc	60
agcccttggc	tccctctgtt	gatcccgccc	cctgtctcag	gcctcactgt	gcaactgctg	120
ctgtcactgc	tgttctctgat	gcctgtccat	ccccagaggt	tggcccggat	gcaggaggat	180
tccccccttg	gaggaggctc	ttctggggaa	ctgacccac	tgggcgagga	ggatctgccc	240
agtgaagagg	attcaccacg	agaggaggat	ccaccgcgag	aggaggatct	acctggagag	300
gaggatctac	ctggagagga	ggatctacct	gaagttaagc	ctaaatcaga	agaagagggc	360
tccctgaagt	tagaggatct	acctactgtt	gaggctcctg	gagatcctca	agaaccccag	420
aataatgcc	acagggacaa	agaaggggat	gaccagagtc	attggcgcta	tggaggcgac	480
ccgcccctgg	cccgggtgtc	cccagcctgc	gcgggcccgt	tccagtcccc	ggtggatatc	540
cgcccccagc	tgcgcgcctt	ctgcccggcc	ctgcgcccc	tggaaactcct	gggcttccag	600
ctcccgcgc	tcccagaact	gcgcctgcgc	aacaatggcc	acagtgtgca	actgaccctg	660
cctcctgggc	tagagatggc	tctgggtccc	gggcgggagt	accgggctct	gcagctgcat	720
ctgcactggg	gggctgcagg	tgcctccggc	tcggagcaca	ctgtggaagg	ccaccgtttc	780
cctgccgaga	tccacgtgg	tcacctcagc	accgcctttg	ccagagttga	cgaggccctg	840
gggcgccccg	gaggcctggc	cgtgtttggc	gcctttcttg	aggaggggcc	ggaagaaaac	900
agtgcctatg	agcagttgct	gtctcgcttg	gaagaaatcg	ctgaggaagg	ctcagagact	960
cagggtcccag	gactggacat	atctgcactc	ctgcctcttg	acttcagccg	ctacttccaa	1020
tatgaggggt	ctctgactac	accgcctgtg	gcccaggggt	tcatctggac	tgtgtttaac	1080
cagracagtga	tgtgtgagtc	taagcagctc	cacaccctct	ctgacaccct	gtggggacct	1140
ggtgactctc	ggctacagct	gaacttccga	gcgacgcagc	ctttgaatgg	gcgagtgtat	1200
gaggcctcct	tccctgctgg	agtggacagc	agtcctcggg	ctgctgagcc	agtccagctg	1260
aatctctgcc	tggctgctgg	tgacatccta	gccctggttt	ttggcctcct	ttttgctgtc	1320
accagcgtcg	cgttccctgt	gcagatgaga	aggcagcaca	gaaggggaac	caaagggggg	1380
gtgagctacc	gcccagcaga	ggtagccgag	actggagcct	agaggctgga	tcttgagaaa	1440
tgtgagaagc	cagccagagg	catctgaggg	ggagccggta	actgtcctgt	cctgctcatt	1500
atgccacttc	cttttaactg	ccaagaaatt	ttttaaaata	aatatttata	at	1552

<210> 66

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001216

<400> 66

tcctgtcctg ctcattatgc cacttccttt taactgccaa gaaatTTTTT aaaataaata 60

<210> 67

<211> 2653

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001254

<400> 67

gagcgcggct	ggagtttgct	gctgccgctg	tgcagtttgt	tcaggggctt	gtgggtggtga	60
gtccgagagg	ctgcgtgtga	gagacgtgag	aaggatcctg	cactgaggag	gtggaaagaa	120
gaggattgct	cgaggaggcc	tggggtctgt	gaggcagcgg	agctgggtga	aggctgcggg	180
ttccggcgag	gcctgagctg	tgctgtcgtc	atgcctcaaa	cccgatccca	ggcacaggct	240
acaatcagtt	ttccaaaaag	gaagctgtct	cgggcattga	acaaagctaa	aaactccagt	300
gatgccaaac	tagaaccaac	aaatgtccaa	accgtaacCt	gttctcctcg	tgtaaaagcc	360
ctgcctctca	gccccaggaa	acgtctgggc	gatgacaaCc	tatgcaacac	tccccattta	420
cctccttggt	ctccaccaaa	gcaaggcaag	aaagagaatg	gtccccctca	ctcacatata	480
cttaagggac	gaagattggt	atttgacaat	cagctgacaa	ttaagtctcc	tagcaaaaga	540
gaactagcca	aagttcacca	aaacaaaata	ctttcttcag	ttagaaaaag	tcaagagatc	600
acaacaaatt	ctgagcagag	atgtccactg	aagaaagaat	ctgcatgtgt	gagactattc	660
aagcaagaag	gcacttgcta	ccagcaagca	aagctgggtCc	tgaacacagc	tgtcccagat	720
cggctgcctg	ccagggaaaag	ggagatggat	gtcatcagga	atthcttgag	ggaacacatc	780
tgtgggaaaa	aagctggaag	cctttacctt	tctgggtgctc	ctggaactgg	aaaaactgcc	840
tgcttaagcc	ggattctgca	agacctcaag	aaggaactga	aaggctttaa	aactatcatg	900
ctgaattgca	tgtccttgag	gactgccag	gctgtattCc	cagctattgc	tcaggagatt	960
tgtcaggaag	aggatccag	gccagctggg	aaggacatga	tgaggaaatt	ggaaaaacat	1020
atgactgcag	agaagggccc	catgattgtg	tgggtattgg	acgagatgga	tcaactggac	1080
agcaaaggcc	aggatgtatt	gtacacgcta	tttgaatggc	catggctaag	caattctcac	1140
ttgggtctga	ttgggtattgc	taataccctg	gatctcacag	atagaattct	acctaggcct	1200
caagctagag	aaaaatgtaa	gccacagctg	ttgaacttCc	cacctatac	cagaaatcag	1260
atagtcacta	ttttgcaaga	tcgacttaat	caggatctta	gagatcaggt	tctggacaat	1320
gctgcagttc	aattctgtgc	ccgcaaagtc	tctgctgttt	caggagatgt	tcgcaaagca	1380
ctggatgttt	gcaggagagc	tattgaaatt	gtagagtcag	atgtcaaaaag	ccagactatt	1440
ctcaaacacc	tgtctgaatg	taaatcacct	tctgagcctc	tgattcccaa	gagggttggt	1500
cttatttcaca	tatcccaagt	catctcagaa	gttgatggta	acaggatgac	cttgagccaa	1560
gaaggagcac	aagattcctt	ccctcttcag	cagaagatCt	tggtttgctc	tttgatgctc	1620
ttgatcaggc	agttgaaaat	caaagaggtc	actctgggga	agttatatga	agcctacagt	1680
aaagtctgtc	gcaaacagca	ggtggcggct	gtggaaccaGt	cagagtgttt	gtcactttca	1740
gggctcttgg	aagccagggg	catttttagga	ttaaagagaa	acaaggaaac	ccgtttgaca	1800
aagggtgttt	tcaagattga	agagaaagaa	atagaacaTg	ctctgaaaga	taaagcttta	1860
attggaaata	tcttagctac	tggattgcct	taaattctTc	tcttacaccc	cacccgaaag	1920

tattcagctg	gcatttagag	agctacagtc	ttcattttag	tgctttacac	attcgggcct	1980
gaaaacaaat	atgacctttt	ttacttgaag	ccaatgaaTt	ttaatctata	gattctttaa	2040
tattagcaca	gaataatatc	tttgggtcct	actattttTa	cccataaaag	tgaccaggta	2100
gacctttttt	aattacattc	actactctta	ccacttgtgt	atctctagcc	aatgtcttg	2160
caagtgtaca	gatctgtgta	gaggaatgtg	tgtatattTa	cctcttcgtt	tgctcaaaca	2220
tgagtgggta	tttttttggt	tgtttttttt	gttggtgtTg	tttttgaggc	gcgtctcacc	2280
ctggtgccca	ggctggagtg	caatggcgcg	ttctctgcTc	actacagcac	ccgcttccca	2340
ggttgaagtg	attctcttgc	ctcagcctcc	cgagtagcTg	ggattacagg	tgcccaccac	2400
cgcgccagc	taatttttta	atthtttagta	gagacagggt	tttaccatgt	tggccaggct	2460
ggtcttgaac	tcctgaccct	caagtgatct	gccacctTg	gcctccctaa	gtgctgggat	2520
tataggcgtg	agccaccatg	ctcagccatt	aaggatttTt	gttaagaact	ttaagtttag	2580
ggtaagaaga	atgaaaatga	tccagaaaaa	tgcaagcaAg	tccacatgga	gatttggagg	2640

acactgggta aag 2653

<210> 68
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001254

<400> 68
 caaggaaacc cgtttgacaa aggtgttttt caagattgaa gagaaagaaa tagaacatgc 60

<210> 69
 <211> 627
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001323

<400> 69
 gcggccgcaa gctcggcact cacggctctg agggctccga cggcactgac ggccatggcg 60
 cgttcgaacc tcccgctggc gctgggcctg gccctggctg cattctgcct cctggcgctg 120
 ccacgcgacg cccggggccg gccgcaggag cgcattggctg gagaactccg ggacctgtcg 180
 cccgacgacc cgcagggtgca gaaggcggcg caggcggccg tggccagcta caacatgggc 240
 agcaacagca tctactactt ccgagacacg cacatcatca aggcgcagag ccagctgggtg 300
 gccggcatca agtacttcct gacgatggag atggggagca cagactgccg caagaccagg 360
 gtcactggag accacgtcga cctcaccact tgccccctgg cagcaggggc gcagcaggag 420
 aagctgcgct gtgactttga ggtccttggt gttccctggc agaactcctc tcagctccta 480
 aagcacaact gtgtgcagat gtgataagtc cccgagggcg aaggccattg ggtttggggc 540
 catggtggag ggcacttcag gtccgtgggc cgtatctgtc acaataaatg gccagtgtcg 600
 ctctcttgcaa aaaaaaaaaa aaaaaaa 627

<210> 70
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001323

<400> 70
 atcaagtact tcctgacgat ggagatgggg agcacagact gccgcaagac cagggtcact 60

<210> 71
 <211> 1812
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001428

<400> 71
 tagctaggca ggaagtcggc gcggggcggcg cggacagtat ctgtgggtac cgggagcacg 60
 gagatctcgc cggcttttac ttcacctcgg tgtctgcagc accctccgct tcctctccta 120
 ggcgacgaga ccagtggtct agaagttcac catgtctatt ctcaagatcc atgccaggga 180
 gatctttgac tctcgcggga atcccactgt tgaggttgat ctcttcacct caaaaggtct 240
 ctctcagagct gctgtgccca gtggtgcttc aactggtatc tatgaggccc tagagctccg 300
 ggacaatgat aagactcgct atatggggaa ggggtgtctc aaggctgttg agcacatcaa 360
 taaaactatt gcgcctgccc tggttagcaa gaaactgaa gtacacagaac aagagaagat 420

tgacaaactg	atgatcgaga	tggatggaac	agaaaataaa	tctaagtttg	gtgcgaacgc	480
cattctgggg	gtgtcccttg	ccgtctgcaa	agctggtgc	ggtgagaagg	gggtccccc	540
gtaccgccac	atcgctgact	tggctggcaa	ctctgaagtc	atcctgccag	tcccggcggt	600
caatgtcatc	aatggcggtt	ctcatgctgg	caacaagctg	gccatgcagg	agttcatgat	660
cctcccagtc	ggtgcagcaa	acttcagggg	agccatgcgc	attggagcag	aggtttacca	720
caacctgaag	aatgtcatca	aggagaaata	tgggaaagat	gccaccaatg	tgggggatga	780
aggcggttt	gctcccaaca	tcctggagaa	taaagaaggc	ctggagctgc	tgaagactgc	840
tattgggaaa	gctggctaca	ctgataagg	ggtcatcggc	atggacgtag	cggcctccga	900
gttcttcagg	tctgggaagt	atgacctgga	cttcaagtct	cccgatgacc	ccagcaggta	960
catctcgcct	gaccagctgg	ctgacctgta	caagtccttc	atcaaggact	acccagtggg	1020
gtctatcgaa	gatccctttg	accaggatga	ctggggagct	tggcagaagt	tcacagccag	1080
tgcaggaatc	caggtagtgg	gggatgatct	cacagtgcac	aacccaaaga	ggatcgccaa	1140
ggcgtgaac	gagaagtcct	gcaactgcct	cctgctcaaa	gtcaaccaga	ttggctccgt	1200
gaccgagtct	cttcaggcgt	gcaagctggc	ccaggccaat	ggttggggcg	tcatggtgtc	1260
tcacgttctg	ggggagactg	aagatacctt	catcgctga	ctggttgtgg	ggctgtgcac	1320
tgggcagatc	aagactgggtg	ccccttgccg	atctgagcgc	ttggccaagt	acaaccagct	1380
cctcagaatt	gaagaggagc	tgggcagcaa	ggctaagtct	gccggcagga	acttcagaaa	1440
ccccttggcc	aagtaagctg	tgggcaggca	agcccttcgg	tcacctgttg	gctacacaga	1500
cccctccctt	cgtgtcagct	caggcagctc	gaggcccccg	accaacactt	gcaggggtcc	1560
ctgctagtta	gcgccccacc	gccgtggagt	tcgtaccgct	tccttagaac	ttctacagaa	1620
gccaagctcc	ctggagccct	gttggcagct	ctagctttgc	agtcgtgtaa	ttggcccaag	1680
tcattgtttt	tctcgctcca	cttccacca	agtgtctaga	gtcatgtgag	cctcgtgtca	1740
tctccggggt	ggccacaggc	tagatccccg	gtggttttgt	gctcaaaata	aaaagcctca	1800
gtgacccatg	ag	1812				

<210> 72

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001428

<400> 72

agctctagct	tttgcagtcg	tgtaatgggc	ccaagtcatt	gtttttctcg	cctcactttc	60
------------	------------	------------	------------	------------	------------	----

<210> 73

<211> 8368

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001456

<400> 73

gcgatccggg	cgccaccccg	cggtcatcgg	tcaccggctg	ctctcaggaa	cagcagcgca	60
acctctgctc	cctgcctcgc	ctcccgcgcg	cctaggtgc	tgcgacttta	attaaagggc	120
cgtcccctcg	ccgaggctgc	agcaccgccc	ccccggcttc	tcgcgcctca	aaatgagtag	180
ctcccactct	cgggcgggcc	agagcgcagc	aggcgcggct	ccgggcggcg	gcgtcgacac	240
gcgggacgcc	gagatgccgg	ccaccgagaa	ggacctggcg	gaggacgcgc	cgtggaagaa	300
gatccagcag	aacactttca	cgcgctgggt	caacgagcac	ctgaagtgcg	tgagcaagcg	360
catcgccaac	ctgcagacgg	acctgagcga	cgggctgcgcg	cttatcgcgc	tgttggagggt	420

gctcagccag	aagaagatgc	accgcaagca	caaccagcgg	cccactttcc	gccaatgca	480
gcttgagaac	gtgtcggtgg	cgctcgagtt	cctggaccgc	gagagcatca	aactggtgtc	540
catcgacagc	aaggccatcg	tggacgggaa	cctgaagctg	atcctggggc	tcactctggac	600
cctgatcctg	cactactcca	tctccatgcc	catgtgggac	gaggaggagg	atgaggaggc	660
caagaagcag	acccccaaagc	agaggctcct	gggctggatc	cagaacaagc	tgccgcagct	720
gcccatcacc	aacttcagcc	gggactggca	gagcggccgg	gccctggggc	ccctggtgga	780
cagctgtgcc	ccgggcctgt	gtcctgactg	ggactcttgg	gacgccagca	agcccgttac	840

caatgcgcga	gaggccatgc	agcaggcgga	tgactggctg	ggcatcccc	aggtgatcac	900
ccccgaggag	attgtggacc	ccaacgtgga	cgagcactct	gtcatgacct	acctgtccca	960
gttccccaa	gccaaactga	agccaggggc	tccttgcgc	cccaaactga	acccgaagaa	1020
agcccgtgcc	tacgggcccag	gcatcgagcc	cacagcgaac	atgggtgaaga	agcgggcaga	1080
gttactctgt	gagaccagaa	gtgctggcca	gggagagggt	ctgggtgtacg	tggaggaccc	1140
ggccggacac	caggaggagg	caaaagtac	cgccaataac	gacaagaacc	gcaccttctc	1200
cgtctggtac	gtccccgagg	tgacggggac	tcataagggt	actgtgctct	ttgctggcca	1260
gcacatcgcc	aagagccct	tcgagggtga	cgtggataag	tcacagggtg	acgccagcaa	1320
agtgcagcc	caagggtccc	gcctggagcc	cagtggcaac	atcgccaaca	agaccaccta	1380
ctttgagatc	tttacggcag	gagctggcac	gggagagggt	gagggttgtga	tccaggaccc	1440
catgggacag	aaggggacgg	tagagcctca	gctggaggcc	cggggcgaca	gcacataacc	1500
ctgcagctac	cagcccacca	tggaggcggt	ccaca ccgtg	cacgtcacgt	ttgccggcgt	1560
gcccacccct	cgcagccct	acactgtcac	tggttgccaa	gcctgtaacc	cgagtgcctg	1620
ccgggcggtt	ggccggggcc	tccagcccaa	gggtg tgcgg	gtgaaggaga	cagctgactt	1680
caagggtgtac	acaaagggcg	ctggcagtg	ggagc tgaag	gtcacccgtga	agggccccaa	1740
gggagaggag	cgcggtgaag	agaaggacct	gggggatggc	gtgtatggct	tcgagtatta	1800
ccccatggtc	cctggaacct	atatcgtcac	catca cgtgg	ggtggtcaga	acatcgggcg	1860
cagtccttc	gaagtgaagg	tgggcaccga	gtgtgccaat	cagaagggtac	gggcctgggg	1920
ccctgggctg	gaggggcgcg	tcgttggcaa	gtcagcagac	tttgtgggtg	aggctatcgg	1980
ggacgacgtg	ggcacgctgg	gcttctcggt	ggaaggcca	tcgcaggcta	agatcgaatg	2040
tgacgacaag	ggcgacggct	cctgtgatgt	gcgct actgg	ccgcaggagg	ctggcgagta	2100
tgccgttcac	gtgctgtgca	acagcgaaga	catcc gcctc	agcccttca	tggtgacat	2160
ccgtgacgcg	ccccaggact	tccaccaga	cagggtgaag	gcacgtgggc	ctggattgga	2220
gaagacaggt	gtggccgtca	acaagccagc	agagt tcaca	gtggatgcca	agcacgggtg	2280
caaggcccca	cttcgggtcc	aagtccagga	caatgaaggc	tgccctgtgg	aggcgttggt	2340
caaggacaac	ggcaatggca	cttacagctg	ctctc actg	cccaggaagc	cgggtgaagca	2400
cacagccatg	gtgtcctggg	gaggcgtag	catcc ccaac	agcccttca	gggtgaatgt	2460
gggagctggc	agccacccca	acaagggtcaa	agtat acggc	cccggagtag	ccaagacagg	2520
gctcaaggcc	cacgagccca	cctacttcac	tgtggactgc	gccgaggctg	gccaggggga	2580
cgtcagcatc	ggcatcaagt	gtgcccctgg	agtgg taggc	cccgccgaag	ctgacatcga	2640
cttcgacatc	atccgcaatg	acaatgacac	cttca cggtc	aagtacacgc	cccggggggc	2700
tggcagctac	accattatgg	tcctctttgc	tgacc agggc	acgccacca	gccccatccg	2760
agtcaagggt	gagccctctc	atgacgccag	taagggtgaag	gccgaggggc	ctggcctcag	2820
tcgcaactgg	gtcgagcttg	gcaagccac	ccact tcaca	gtaaatgcca	aagctgctgg	2880
caaaggcaag	ctggacgtcc	agttctcagg	actca.ccaag	ggggatgcag	tgcgagtggt	2940
ggacatcatc	gaccaccatg	acaacaccta	cacagrtcaag	tacacgcctg	tccagcaggg	3000
tccagtaggc	gtcaatgtca	cttatggagg	ggatccatc	cctaagagcc	ctttctcagt	3060
ggcagtatct	ccaagcctgg	acctcagcaa	gatca aggtg	tctggcctgg	gagagaagggt	3120
ggacgttggc	aaagaccagg	agttcacagt	caaat caaag	ggtgctgggt	gtcaaggcaa	3180
agtggcatcc	aagattgtgg	gcccctcggg	tgcagcgggt	ccctgcaagg	tggagccagg	3240
cctgggggct	gacaacagt	tggtgcgctt	cctgc cccgt	gaggaagggc	cctatgaggt	3300
ggaggtgacc	tatgacggcg	tgcccgtgcc	tggcagcccc	tttctctgg	aagctgtggc	3360
ccccaccag	cctagcaagg	tgaaggcgtt	tgggc cgggg	ctgcaggggg	gcagtcgggg	3420
ctcccccgcc	cgcttcacca	tcgacaccaa	gggcgcgggc	acagggtggc	tgggcctgac	3480
ggtggagggc	ccctgtgagg	cgcagctcga	gtgct tggac	aatggggatg	gcacatgttc	3540
cgtgtcctac	gtgccaccgc	agcccgggga	ctaca acatc	aacatcctct	tcgctgacac	3600
ccacatccct	ggctcccat	tcaaggccca	cgtgggttccc	tgctttgacg	catccaaagt	3660
caagtgtctc	ggccccgggc	tggagcgggc	caccgctggg	gagggtgggc	aattccaaagt	3720
ggactgctcg	agcgcgggca	gcgcggagct	gaccaattgag	atctgctcgg	aggcggggct	3780
tccggccgag	gtgtacatcc	aggaccacgg	tgatggcacg	cacaccatta	cctacattcc	3840
cctctgcccc	ggggcctaca	ccgtcacct	caagt acggc	ggccagcccg	tgcccaactt	3900
ccccagcaag	ctgcaggtgg	aacctgcggt	ggaca.cttcc	ggtgtccagt	tgctacggcc	3960
tggatttgag	ggcgagggtg	tcttcgctga	ggccaccact	gagttcagtg	tggacggccg	4020
ggctctgaca	cagaccggag	ggccgcacgt	caaggccccgt	gtggccaacc	cctcaggcaa	4080
cctgacggag	acctacgttc	aggaccgtgg	cgatggcatg	tacaaagtgg	agtacacgcc	4140
ttacgaggag	ggactgcact	ccgtggacgt	gacctatgac	ggcagtcctg	tgccagcag	4200
ccccttccag	gtgcccgtga	ccgagggtg	cgaccctcc	cgggtgcgtg	tcacggggcc	4260
aggcatccaa	agtggcacca	ccaacaagcc	caaca agttc	actgtggaga	ccagggggagc	4320
tggcacgggc	ggcctggggc	tggtgttaga	gggc cctcc	gaggccaaga	tgtcctgcat	4380
ggataacaag	gacggcagct	gctcggctga	gtacatccct	tatgaggctg	gcacctacag	4440
cctcaacgtc	acctatgggtg	gccatcaagt	gccaggcag	cctttcaagg	tcctgtgca	4500

t gatgtgaca	gatgcgtcca	aggtcaagtg	ctctgggccc	ggcctgagcc	caggcatggt	4560
t cgtgccaaac	ctccctcagt	ccttccaggt	ggacacaagc	aaggctggtg	tggccccatt	4620
gcaggtcaaa	gtgcaagggc	ccaaaggcct	ggtggagcca	gtggacgtgg	tagacaacgc	4680
t gatggcacc	cagaccgtca	attatgtgcc	cagccgaaga	gggccctaca	gcattctcagt	4740
a ctgtatgga	gatgaagagg	taccccgagg	ccccttcaag	gtcaaggtgc	tgcctactca	4800
t gatgccagc	aaggtgaagg	ccagtggccc	cgggctcaac	accactggcg	tgcctgccag	4860
c ctgcccgtg	gagttcacca	tcgatgcaaa	ggacgcggg	gagggcctgc	tggctgtcca	4920
g atcacggat	cccgaaggca	agccgaagaa	gacacacatc	caagacaacc	atgacggcac	4980
g tatacagtg	gcctacgtgc	cagacgtgac	aggtcgctac	accatcctca	tcaagtacgg	5040
t ggtgacgag	atcccccttct	ccccgtaccg	cgtgcgtgcc	gtgcccaccg	gggacgccag	5100
caagtgcact	gtcacagtgt	caatcggagg	tcacgggcta	ggtgctggca	tcggccccac	5160
c attcagatt	ggggaggaga	cggatgatcac	tgtggacact	aaggcggcag	gcaaaggcaa	5220
a gtgacgtgc	accgtgtgca	cgcctgatgg	ctcagaggtg	gatgtggacg	tgggtggagaa	5280
t gaggacggc	acttttcgaca	tcttctacac	ggccccccag	ccgggcaaat	acgtcatctg	5340
t gtgcgcttt	ggtggcgagc	acgtgcccaa	cagccccctt	caagtgcagg	ctctggctgg	5400
g gaccagccc	tcggtgcagc	ccccctctacg	gtctcagcag	ctggccccac	agtacacct	5460
c gcccagggc	ggccagcaga	cttgggcccc	ggagagggcc	ctggtgggtg	tcaatgggct	5520
g gatgtgacc	agcctgaggc	cctttgacct	tgtcatcccc	ttcaccatca	agaaggggcg	5580
g atcacaggg	gaggttcgga	tgccttcagg	caaggtggcg	cagccccacca	tactgacaa	5640
caaagacggc	accgtgaccg	tgcggtatgc	accagcgag	gctggcctgc	acgagatgga	5700
c atccgctat	gacaacatgc	acatcccagg	aagccccctg	cagttctatg	tggattacgt	5760
caactgtggc	catgtcactg	cctatggggc	tggcctcacc	catggagtag	tgaacaagcc	5820
t gccaccttc	accgtcaaca	ccaaggatgc	aggagagggg	ggcctgtctc	tggccattga	5880
g ggcccgtcc	aaagcagaaa	tcagctgcac	tgacaaacag	gatgggacat	gcagcgtgtc	5940
c tacctgcct	gtgctgccgg	gggactacag	cattctagtc	aagtacaatg	aacagcacgt	6000
c cagggcagc	cccttcactg	ctcgggtcac	aggtgacgac	tccatgcgta	tgtcccacct	6060
aaaggtcggc	tctgctcggc	acatccccat	caacatctca	gagacggatc	tcagcctgct	6120
g acggccact	gtgggtccgc	cctcggggcg	ggaggagccc	tgtttgctga	agcggctgcg	6180
t aatggccac	gtgggggattt	cattcgtgcc	caaggagacg	ggggagcacc	tgggtgcatgt	6240
gaagaaaaat	ggccagcacg	tggccagcag	ccccatcccg	gtggtgatca	gccagtcgga	6300
a attggggat	gccagtcgtg	ttcgggtctc	tggtcagggg	cttcacgaag	gccacacctt	6360
t gagcctgca	gagtttatca	ttgatacccg	cgatgcaggc	tatggtgggc	tcagcctgtc	6420
c attgagggc	cccagcaagg	tggacatcaa	cacagaggac	ctggaggacg	ggacgtgcag	6480
g gtcacctac	tgccccacag	agccaggcaa	ctacatcctc	aacatcaagt	ttgccgacca	6540
g cacgtgcct	ggcagccccct	tctctgtgaa	ggtgacaggc	gagggccggg	tgaaagagag	6600
c atcacccgc	aggcgtcggg	ctccttcagt	ggccaaagct	ggtagtcatt	gtgacctcag	6660
c ctgaaaatc	cctgaaatta	gcattccagga	tatgacagcc	caggtgacca	gccccatcgg	6720
caagacccat	gaggccgaga	tcgtggaagg	ggagaaacac	acctactgca	tcgcctttgt	6780
t cccgctgag	atgggcacac	acacagtcag	cgtcaagtac	aagggccagc	acgtgcctgg	6840
g agccccctc	cagttcacccg	tggggccccct	aggggaaggg	ggagcccaca	aggtccgagc	6900
t gggggccccct	ggcctggaga	gagctgaagc	tggagtggca	gccgaattca	gtatctggac	6960
c cggggaagct	ggtgctggag	gcctggccat	tgtgtcgag	ggccccagca	aggctgagat	7020
c tctttttgag	gaccgcaagg	acggctcctg	tgtgtggct	tatgtggtcc	aggagccagg	7080
t gactacgaa	gtctcagtca	agttcaacga	ggaaacatt	cccacagcc	ccttcgtggt	7140
g cctgtggct	tctccgtctg	gcgacgcccg	ccgcctcact	gtttctagcc	ttcaggagtc	7200
a gggctaaag	gtcaaccagc	cagcctcttt	tgcagt cagc	ctgaacgggg	ccaagggggc	7260
g atcgatgcc	aaggtgcaca	gccccctcagg	agccctggag	gagtgtatg	tcacagaaat	7320
t gaccaagat	aagtatgctg	tgcgcttcat	ccctcgggag	aatggcgttt	acctgattga	7380
c gtcaagttc	aacgttacc	acatccccctg	aagccccttc	aagatccgag	ttggggagcc	7440
t gggcatgga	ggggacccag	gcttgggtgtc	tgttaccgga	gcaggtctgg	aaggcggtgt	7500
c acagggaac	ccagctgagt	tcgtcgtgaa	cacgagcaat	gcgggagctg	gtgccctgtc	7560
g gtgaccatt	gacggccccct	ccaaggtgaa	gatggaatgc	caggagtgc	ctgagggcta	7620
c cgcgtccacc	tataccccca	tggcaacctg	cagcta cctc	atctccatca	agtacggcg	7680
c cccctaccac	attgggggca	gcccccttcaa	ggccaaagtc	acaggcccc	gtctcgtcag	7740
caaccacagc	ctccacgaga	catcatcagt	gtttgt agac	tctctgacca	aggccacctg	7800
t gccccccag	catggggccc	cgggtcctgg	gcctgc tgac	gccagcaagg	tgggtggcaa	7860
g ggcctgggg	ctgagcaagg	cctacgtagg	ccagaaagagc	agcttcacag	tagactgcag	7920
caaagcaggc	aacaacatgc	tgtggtggg	ggttca tggc	ccaaggaccc	cctgcgagga	7980
g atcctggtg	aagcacgtgg	gcagccggct	ctacag cgtg	tcctacctgc	tcaaggacaa	8040
g gggggagtac	acactggtgg	tcaaattgggg	gcacga gcac	atcccaggca	gccccctaccg	8100
c gttgtggtg	ccctgagtc	ggggccccgtg	ccagcc ggca	gcccccaagc	ctgccccgct	8160

```

acccaagcag ccccgccctc ttccctcaaa ccccgGCCCA ggccgccttg gccgcccgcc 8220
tgtcactgca gctgcccctg cctgtgccc tgctgcgctc acctgcctcc ccagccagcc 8280
gctgacctct cggctttcac ttgggcagag ggagccattt ggtggcgctg cttgtcttct 8340
ttggttcttg gagggtgag ggatgggg 8368

```

<210> 74
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001456

<400> 74
 tgacctctcg gctttcactt gggcagaggg agccattttgg tggcgctgct tgtcttcttt 60

<210> 75
 <211> 1642
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 75

```

ccagatctca gaggagcctg gctaagcaaa accctgcaga acggctgcct aatttacagc 60
aaccatgagt acaaatggtg atgatcatca ggtcaaggat agtctggagc aattgagatg 120
tcactttaca tgggagttat ccattgatga cgatgaaatg cctgatttag aaaacagagt 180
cttgatcagc attgaattcc tagacaccaa atacagtgtg ggaatacaca acctactagc 240
ctatgtgaaa cacctgaaa ggcagaatga ggaagcctg aagagcttaa aagaagctga 300
aaactttaatg caggaagaac atgacaacca agcaaatgtg aggagtctgg tgacctggg 360
caactttgcc tggatgtatt accacatggg cagactggca gaagcccaga cttacctgga 420
caaggtggag aacatttgca agaagctttc aaatccttc cgctatagaa tggagtgtcc 480
agaaatagac tgtgaggaag gatgggcctt gctgaagtgt ggaggaaaga attatgaacg 540
ggccaaggcc tgctttgaaa aggtgcttga agtggaccct gaaaaccctg aatccagcgc 600
tgggtatgcg atctctgcct atcgccctgga tggctttaa ttagccacaa aaaatcacia 660
gccattttct ttgcttccc taaggcaggg tgccgctta aatccagaca atggatatat 720
taaggttctc cttgccctga agcttcagga tgaaggacag gaagctgaag gagaaaagta 780
cattgaagaa gctctagcca acatgtctc acagacctat gtctttcgat atgcagccaa 840
gttttaccga agaaaaggct ctgtggataa agctcttgag ttattaaaaa aggccttgca 900
ggaaacaccc acttctgtct tactgcatca ccagataggg ctttgctaca aggcacaaat 960
gatccaaatc aaggaggcta caaaagggca gcctagaggg cagaacagag aaaagctaga 1020
caaaatgata agatcagcca tatttcattt tgaatctgca gtggaaaaaa agccacatt 1080
tgaggtggct catctagacc tggcaagaat gtatatagaa gcaggcaatc acagaaaagc 1140
tgaagagaat tttcaaaaat tgttatgcat gaaaccagtg gtagaagaaa caatgcaaga 1200
catacatttc tactatgggc ggtttcagga atttcaaaag aaatctgacg tcaatgcaat 1260
tatccattat ttaaaagcta taaaaataga acaggcatca ttaacaaggg ataaaagtat 1320
caattctttg aagaaattgg ttttaaggaa acttcggaga aaggcattag atctggaaag 1380
cttgagcctc cttgggttcg tctataaatt ggaaggaaat atgaatgaag cctggagta 1440
ctatgagcgg gccctgagac tggctgctga ctttgaagaa tctgtgagac aaggtoctta 1500
ggcaccacga tatcagccac tttcacattt catttcattt tatgctaaca tttactaatc 1560
atcttttctg cttactgttt tcagaaacat tataatcac tgtaatgatg taattcttga 1620
ataataaatc tgacaaaata tt 1642

```

<210> 76
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 76

gtatcaattc tttgaagaaa ttgggttttaa ggaaacttcg gagaaaggca ttagatctgg 60

<210> 77

<211> 3344

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001605

<400> 77

```

ggtacagctg cgcgtctgcg ggaatagggtg cagcggggccc ttggcggggg actctgaggg 60
aggagctggg gacggcgacc ctaggagagt tctttggggg gactttcaag atggactcta 120
ctctaacagc aagtgaatc cggcagcgat ttatagattt cttcaagagg aacgagcata 180
cgtatgttca ctgctctgcc accatcccat tggatgaccc cactttgctc ttggccaatg 240
caggcatgaa ccagtttaaa cccattttcc tgaacacaat tgacctatct caccatcatg 300
caaagctgag cagagctgcc aatacccaga agtgcatccg ggctgggggc aaacaaaatg 360
acctggacga tgtgggcaag gatgtctatc atcacacctt cttcgagatg ctgggctctt 420
ggtcttttgg agattacttt aaggaattgg catgtaagat ggctctggaa ctctcaccc 480
aagagtttgg cattcccatt gaaagacttt atgttaactta ctttggcggg gatgaagcag 540
ctggcttaga agcagatctg gaatgcaaac agatctggca aaatttgggg ctggatgaca 600
ccaaaatcct cccaggcaac atgaaggata acttctggga gatgggtgac acgggccct 660
gtggctcttg cagtgcagac cactacgacc ggattggtgg tcgggacgcc gcacatcttg 720
tcaaccagga cgacccta atgtctggaga tctggaacct tgtgttcac cagtataaca 780
gggaagctga tggcattctg aaacctcttc ccaagaaaag cattgacaca gggatggggc 840
tggaaacgact ggtatctgtg ctgcagaata agatgtccaa ctatgacact gacctttttg 900
tcccttactt tgaagccatt cagaagggca cagggtgccc accatacact gggaaagtgt 960
gtgctgagga tgccgatggg attgacatgg cctaccgggt gctggctgac catgctcgga 1020
ccatcactgt ggcactgggt gatgggtggc ggctgacaa cacagggcgt ggatatgtgt 1080
tgagacggat tctccgccga gctgtccgat acgcccatga aaagctcaat gccagcaggg 1140
gcttctttgc tacgttagtg gatgtgtgct tccagtccct gggagatgca tttcctgagc 1200
tgaagaagga cccagacatg gtgaaggaca tcattaatga agaagaggtg cagtttctca 1260
agactctcag cagagggcgt cgcacctgtg acaggaaaat tcagagcctg ggagacagca 1320
agaccattcc cggagacact gcttggctcc tctatgacac ctatgggttt ccagtggatc 1380
tgactggact gattgctgaa gagaagggcc tgggtggtaga catggatggc tttgaagagg 1440
agaggaaact ggcccagctg aaatcacagg gcaaggagagc tgggtgggaa gacctatta 1500
tgctggacat ttacgctatc gaagagctcc gggcacgggg tctggagggtc acagatgatt 1560
ccccaaagta caattaccat ttggactcca gtggtagcta tgtatttgag aacacagtgg 1620
ctacggtgat ggctctgcgc agggagaaga tgttcgtgga agaggtgtcc acaggccagg 1680
agtgtggagt ggtgctggac aagacctgtt tctatgctga gcaaggaggc cagatctatg 1740
acgaaggcta cctggtgaag gtggatgaca cagtggaaga taaaacagag tttacagtga 1800
agaatgctca ggtccgagga gggatgtgct tacacatttg aaccatctac ggtgacctga 1860
aagtggggga tcagggtctg ctgtttattg atgagccccg acgaagacc atcatgagca 1920
accacacagc taogcacatt ctgaacttcg cctgcgctc agtgcttggg gaagctgacc 1980
agaaaggctc attggttgct cctgaccgcc tcagatttga ctttactgcc aaggagacca 2040
tgtccacca acagatcaag aaggctgaag agattgctaa tgagatgatt gaggcagcca 2100
aggcctcta taccaggat tgccccctgg cagcagcgaa agccatccag ggcctacggg 2160
ctgtgtttga tgagacctat cctgacctg tgcagctcgt ctccattggg gtcccgtgt 2220
ccgagttgct ggtgacccc tctgggctcg ctggctcct gacttctgtt gatttctgtg 2280
ggggaacgca cctgcggaac tctgagtcag caggagcttt tgtgatcgtg acggaagaag 2340
ccattgccaa ggggtatccg aggtatgtgg ctgtcacagg tgccgaggcc cagaaggccc 2400
tcaggaaagc agagagcttg aagaaatgtc tctctgtcat ggaagccaaa gtgaaggctc 2460
agactgctcc aaacaaggat gtgcagaggg agatcgtgta ccttgagag gccctggcca 2520
ctgcagtcac ccccagtgga cagaaggatg aattgcggga gactctcaa tccctaaaga 2580
aggtcatgga tgacttggac cgagccagca aagccgatgt ccagaaacga gtgttagaga 2640

agacgaagca gttcatcgac agcaacccca accagcctct tgtcatcctg gagatggaga 2700
gcggcgccctc agccaaggcc ctgaatgaag ccttgaagct cttcaagatg cactcccctc 2760

```

```

agactttctgc catgctcttc acggtgggaca atgaggtctgg caagatcacg tgcctgtgtc 2820
aagtccccca gaatgcagcc aatcggggct taaaagccag cgagtgggtg cagcaggtgt 2880
caggcttgat ggacggtaaa ggtggtggca aggatgtgtc tgcacaggcc acaggcaaga 2940
acgttggtg cctgcaggag gcgtgcagc tggccacttc cttcgcccag ctgcgcctcg 3000
gggatgtaaa gaactgagtg gggaaggagg aggctccac tggatccatc cgtccagcca 3060
agagctcttc atctgctaca agaacatttg aatcttggga cctttaaaga gcccctccta 3120
accagcagt aactggaaca cacttgggag cagtcctatg tctcagtgcc ccttaaattt 3180
ctgccctgag ccctccacgt cagtgccatc ggtctagaac cactaaccct gcattgctgt 3240
tgatcgtcac gctcgcatct atagataacg gctctccaga cctgagcttt ccgcgtcagc 3300
aagtaggaat cgtttttctg gcagagaata aaaggaccac gtgc 3344

```

```

<210> 78
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001605

```

```

<400> 78
gccaagagct cttcatctgc tacaagaaca tttagaatctt gggaccttta aagagcccct 60

```

```

<210> 79
<211> 417
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001645

```

```

<400> 79
acctcccaac caagccctcc agcaaggatt caggagtgcc cctcgggcct cgccatgagg 60
ctcttctctgt cgctcccgtt cctgggtggtg gttctgtcga tcgtcttgga aggccagacc 120
ccagcccagg ggaccccaga cgtctccagt gccttggata agctgaagga gtttggaaac 180
aacttgagg acaaggctcg ggaactcatc agccgcatca aacagagtga actttctgcc 240
aagatgcggg agtggttttc agagacattt cagaaagtga aggagaaact caagattgac 300
tcatgaggac ctgaagggtg acatccagga ggggcctctg aaatttccca cccccagcg 360
cctgtgctga ggactccgc catgtggccc caggtgccac caataaaaat cctaccg 417

```

```

<210> 80
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001645

```

```

<400> 80
aaacagagtg aactttctgc caagatgcgg gagtgggttt cagagacatt tcagaaagtg 60

```

```

<210> 81
<211> 1389
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001809

```

```

<400> 81
cgcggacttc tgccaagcac cggctcatgt gaggctcgcg gcacagcgtt ctctgggctc 60

```



```

cccagaagcc agccttttcg tcccggaccc ggcagcccca gcaggagccg tgggaccggg 120
cgccagcacc ctctgcggcg tgtcatgggc ccgcgccgcc ggagccgaaa gcccagggcc 180
ccgaggaggg gcagcccag cccgaccccc accCccggcc cctcccggcg gggcccctcc 240
ttaggcgctt cctcccatca acacagtcgg cggagacaag gttgggctaaa ggagatccga 300
aagcttcaga agagcacaca cctcttgata aggaagctgc ccttcagccg cctggcaaga 360
gaaatatgtg ttaaattcac tcgtggtgtg gacttcaatt ggcaagccca ggcctattg 420
gccctacaag aggcagcaga agcatttcta gttcatctct ttgaggacgc ctatctctc 480
accttacatg caggccgagt tactctcttc ccaaaggatg tgcaactggc ccggaggatc 540
cggggccttg aggagggact cggctgagct cctgcaccca gtgtttctgt cagtctttcc 600
tgctcagcca ggggggatga taccggggac tctccagagc catgactaga tccaatggat 660
tctgcgatgc tgtctggact ttgctgtctc tgaacagtat gtgtgtgttg ctttaaatat 720
ttttcttttt tttgagaagg agaagactgc atgactttcc tctgtaacag aggtaatata 780
tgagacaatc aacaccgttc caaaggcctg aaaataattt tcagataaag agactccaag 840
gttgacttta gtttgtgagt tactcatgtg actatttgag gattttgaaa acatcagatt 900
tgctgtggta tgggagaaaa ggttatgtac ttaattattt agctctttct gtaatattta 960
cattttttac catatgtaca tttgtacttt tattttacac ataagggaaa aaataagacc 1020
actttgagca gttgcctgga aggcctggga tttccatcat atagacctct gcccttcaga 1080
tagcctcac cattagtggc agcatcatgt aactgagtggt actgtgcttg tcaacggatg 1140
tgtagctttt cagaaaacta attggggatg aatagaaaac ctgtaagctt tgatgttctg 1200
gttacttcta gtaaattcct gtcaaaatca attcagaaat tctaacttgg agaatttaac 1260
attttactct tgtaaatacat agaagatgta tcataacagt tcagaatttt aaagtacatt 1320
ttcgatgctt ttatgggtat tttttagatt tctttgtaga gagataataa aaatcaaaat 1380
atttaatga 1389

```

<210> 82

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001809

<400> 82

```

ggggatgaat agaaaacctg taagctttga tgtctctgggt acttctagta aattcctgtc 60

```

<210> 83

<211> 2205

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001909

<400> 83

```

gcgcacgcgc gccgcgcca cgtgaccggt ccgggtgcaa acacgcgggt cagctgatcc 60
ggcccaactg cggcgtcatc ccggctataa gcgcacggcc tcggcgaccc tctccgaccc 120
ggccgcgcgc gccatgcagc cctccagcct tctgccgctc gccctctgcc tgctggctgc 180
accgcctcc gcgctcgtca ggatcccgtt gcaCaagttc acgtccatcc gccggaccat 240
gtcggagggt gggggctctg tggaggacct gatTgccaaa ggccccgtct caaagtactc 300
ccaggcgggt ccagccgtga ccgagggggc catTccccgag gtgctcaaga actacatgga 360
cgcccagtac tacggggaga ttggcatcgg gacgcccccc cagtgtctca cagtcgtctt 420
cgacacgggc tctccaacc tgtgggtccc ctccatccac tgcaaactgc tggacatcgc 480
ttgctggatc caccacaagt acaacagcga caagtccagc acctacgtga agaattgtac 540
ctcgtttgac atccactatg gctcgggcag cctctccggg tacctgagcc aggacactgt 600
gtcgggtgcc tgccagtcag cgtcgtcagc ctctgccctg ggcggtgtca aagtggagag 660
gcaggtcttt ggggaggcca ccaagcagcc aggcataacc ttcatcgag ccaagttcga 720
tggcatcctg ggcattggcct acccccgcct ctccgtcaac aacgtgctgc ccgtcttcga 780
caacctgatg cagcagaagc tgggtggacca gaaCatcttc tccttctacc tgagcaggga 840
cccagatgcg cagcctgggg gtgagctgat gctgggtggc acagactcca agtattacaa 900

```

```

gggtttctctg tccctacctga atgtcacccg caaggcctac tggcaggtcc acctggacca 960
ggtggaggtg gccagcgggc tgaccctgtg caaggagggc tgtgaggcca ttgtggacac 1020
aggcacttcc ctcatggttg gcccggtgga tgaggtgcgc gagctgcaga aggccatcgg 1080
ggccgtgccg ctgattcagg gcgagtacat gatccccctgt gagaaggtgt ccaccctgcc 1140
cgcgatcaca ctgaagctgg gaggcaaagg ctacaagctg tccccagagg actacacgct 1200
caaggtgtcg caggccggga agaccctctg cctgagcggc ttcattgggca tggacatccc 1260
gccaccagc gggccactct ggatcctggg cgacgtcttc atcgccgct actacactgt 1320
gtttgaccgt gacaacaaca ggtggggctt cggcagggct gccgcctct agttcccaag 1380
cggtccgcgc gccagcacag aaacagagga gagtcccaga gcaggaggcc cctggcccag 1440
cggccccctc cacacacacc cacacactcg cccgcccact gtcttgggcg ccctggaagc 1500
cgggggccca agcccgaactt gctgttttgt tctgtgtgtt tcccctccct ggggttcagaa 1560
atgtgcctg cctgtctgtc tctccatctg ttggtgggg gtagagctga tccagagcac 1620
agatctgtt cgtgcattgg aagacccac ccaagcttgg cagccgagct cgtgtatcct 1680
ggggctccct tcatctccag ggagtccct ccccgccct accagcgccc gctgggctga 1740

```

```

gcccctaccc cacaccaggc cgtcctccc gggccctccc tggaaacctg ccctgcctga 1800
gggcccctct gccagcttg gcccagctg ggctctgcca ccctacctgt tcagtgtccc 1860
gggcccgttg aggatgaggc cgctagaggc ctgaggatga gctggaagga gtgagagggg 1920
acaaaacca ccttgttgga gcctgcaggg tgggtgctggg actgagccag tcccaggggc 1980
atgtattggc ctggaggtgg ggttgggatt gggggctggt gccagccttc ctctgcagct 2040
gacctctgtt gtccctccc tgggcggctg agagccccag ctgacatgga aatacagttg 2100
ttggcctccg gcctcccctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2160
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa 2205

```

<210> 84
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001909

```

<400> 84
tctgtttggt ggggtagag ctgatccaga gcacagatct gtttcgtgca ttggaagacc 60

```

<210> 85
 <211> 817
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002038

```

<400> 85
gaacggttta ctgctgctg tgccatcta tcagcaggct ccgggctgaa gattgcttct 60
cttctctcct ccaaggtcta gtgacggagc ccgcgcgcgg gccaccatg cggcagaagg 120
cggatctcgt tttcttgtgc tacctgctgc tcttcacttg cagtggggtg gaggcaggta 180
agaaaaagtg ctcgagagc tcggacagcg gctccgggtt ctggaaggcc ctgaccttca 240
tggccgtcgg aggaggact gcagtcgccg ggttgccgc gctgggcttc accggcgccg 300
gcctcgccgc caactcggtg gctgcctcgc tgatgagctg gtctgcgata ctgaatgggg 360
gcggcgctgcc cgccgggggg ctagtggcca cgctgcagag cctcggggct ggtggcagca 420
gcgtcgatcat aggtaatatt ggtgccctga tgggctacgc caccacaag tatctcgata 480
gtgaggagga tgaggagtag ccagcagctc ccagaacctc ttcttcttct ttggcctaac 540
tcttccagtt aggatctaga actttgcctt tttttttttt tttttttttt tttgagatgg 600
gtttctacta tattgtccag gctagagtgc agtggtatt cacagatgcg aacatagtag 660
actgcagcct ccaactccta gcctcaagtg atcctcctgt ctcaacctcc caagtaggat 720
tacaagcatg cgccgacgat gccagaatc cagaactttg tctatcactc tcccacaaca 780
cctagatgtg aaaacagaat aaacttcacc cagaaaa 817

```

<210> 86
 <211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_002038

<400> 86
agctcccaga acctcttctt ccttcttggc ctaactcttc cagttaggat ctagaacttt 60

<210> 87
<211> 1283
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 87
ctctctgctc ctctgttctg acagtcagc c gcatcttctt ttgcgtcgcc agccgagcca 60
catcgctcag acaccatggg gaaggtgaa g gtcggagtca acggatttgg tcgtattggg 120
cgcttggtca ccagggctgc ttttaactc t ggtaaagtgg atattgttgc catcaatgac 180
cccttcattg acctcaacta catggttta c atgttccaat atgattccac ccattggcaa 240
ttccatggca ccgtcaaggc tgagaacggg aagcttgtca tcaatggaaa tcccatcacc 300
atcttccagg agcgagatcc ctccaaaat c aagtggggcg atgctggcg tgagtacgtc 360
gtggagtcca ctggcgtctt caccaccat g gagaaggctg gggctcattt gcagggggga 420
gccaaaaggg tcatcatctc tgccccctc t gctgatgcc ccattgtcgt catgggtgtg 480
aaccatgaga agtatgacaa cagcctcaa g atcatcagca atgcctcctg caccaccaac 540
tgcttagcac ccctggccaa ggtcatcca t gacaactttg gtatcgtgga aggactcatg 600
accacagtcc atgccatcac tgccacca g aagactgtgg atggcccctc cgggaaactg 660
tggtcgtgatg gccgcggggc tctccagaa c atcatccctg cctctactgg cgctgccaa 720
gctgtgggca aggtcatccc tgagctgaa c gggaagctca ctggcatggc cttccgtgtc 780
cccactgcc aactgtcagt ggtggacct g acctgccgtc tagaaaaacc tgccaaatat 840
gatgacatca agaaggtggg gaagcaggc g tcggagggcc ccctcaaggg catcctgggc 900
tacactgagc accaggtggg ctctctga c ttcaacagc acaccactc ctccacctt 960
gacgtgggg ctggcattgc cctcaacga c cactttgtca agctcatttc ctggatatgac 1020
aacgaatttg gctacagcaa caggggtgg t gacctcatg cccacatggc ctccaaggag 1080
taagaccct ggaccaccag cccagcaa g agcacaagag gaagagagag accctcactg 1140
ctggggagtc cctgccacac tcagtcccc c accacactga atctcccctc ctacagttg 1200
ccatgtagac cccttgaaga ggggagggg c taggggagcc gcaccttgtc atgtaccatc 1260
aataaagtac cctgtgctca acc 1283

<210> 88
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002046

<400> 88
ctcaacgacc actttgtcaa gctcatttc c ttggtatgaca acgaatttgg ctacagcaac 60

<210> 89
<211> 1610
<212> DNA
<213> Homo sapiens

<300>

<308> NM_002061

<400> 89

```

ggcacgagggc tgcggcccgca gtagccggag cccgagccgc agccaccggt gccttccttt 60
cccgccgccc cccagccgcc gtccggcctc cctcggggccc gagcgccagac caggctccag 120
ccgcgcggcg ccggcagcct cgcgctccc tctcgggtctc tctcgggcct cgggcaccgc 180
gtcctgtggg cggccgcctg cctgcccgcc cgcccgcagc cccttgcttg ccggcccttg 240
ggcgggcccg gccatgggca ccgacagccg cgcggccaag gcgctcctgg cgcggggccc 300
caccctgcac ctgcagacgg ggaacctgct gaactggggc cgctgcgga agaagtgcc 360
gtccacgcac agcgaggagc ttcattgatt tatccaaaa accttgaatg aatggagttc 420
ccaaatcaac ccagatttgg tcaggaggtt tccagatgtc ttggaatgca ctgtatctca 480
tgcatgtaga aagataaatc ctgatgaaag agaagaaatg aaagtctctg caaaactgtt 540
cattgtagaa tcaaactcct catcatcaac tagaagtgca gttgacatgg cctgttcagt 600

```

```

ccttggagtt gcacagctgg attctgtgat cattgcttca cctcctattg aagatggagt 660
taatctttcc ttggagcatt tacagcctta ctgggaggaa ttagaaaact tagttcagag 720
caaaaagatt gttgccatag gtacctctga tctagacaaa acacagttgg aacagctgta 780
tcagtgggca caggtaaaac caaatagtaa ccaagttaat cttgcctcct gctgtgtgat 840
gccaccagat ttgactgcat ttgctaaaca atttgacata cagctgttga ctcaaatga 900
tccaaaagaa ctgctttctg aagcaagttt ccaagaagct cttcaggaaa gcattcctga 960
cattcaagcg cagcagtggtg tgccgctgtg gctactgcgg tattcgggtc ttgtgaaaag 1020
tagaggaatt atcaaatcaa aaggctacat tttacaagct aaaagaaggg gttcttaact 1080
gacttaggag cataacttac ctgtaatttc cttcaatatg agagaaaatt gagatgtgta 1140
aaatctagtt actgcctgta aatgggtgtc ttgaggcaga tattctttcg tcatatttga 1200
cagtatgttg tctgtcaagt tttaaatact tatcttgctt ccatatcaat ccattctcat 1260
gaacctctgt attgctttcc ttaaactatt gttttctaat tgaaattgtc tataaagaaa 1320
atacttgcaa tatatttttc ctttattttt atgactaata taaatcaaga aaatttggtg 1380
ttagatatat tttggcctag gtatcagggt aatgtatata catatttttt atttccaaaa 1440
aaaattcatt aattgcttct taactcttat tataaccaag caatttaatt acaattgtta 1500
aaactgaaat actggaagaa gatatttttc ctgtcattga tgagatatat cagagtaact 1560
ggagtagctg ggatttacta gtagtgtaaa taaaattcac tcttcaatac 1610

```

<210> 90

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002061

<400> 90

```

ctgacttagg agcataactt acctgtaatt tccttcaata tgagagaaaa ttgagatgtg 60

```

<210> 91

<211> 873

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002106

<400> 91

```

cgcagtttga atcgcggtgc gacgaaggag taggtggtgg gatctcaccg tgggtccgat 60
tagccttttc tctgccttgc ttgcttgagc ttcagcggaa ttcgaaatgg ctggcggtaa 120
ggctggaaag gactccggaa aggccaagac aaaggcgggt tcccgctcgc agagagccgg 180
cttgcaagtt ccagtgggca gtattcatcg acacctaaaa tctaggacga ccagtcattg 240
acgtgtgggc gcgactgccg ctgtgtacag cgcagccatc ctggagtacc tcaccgcaga 300
ggtacttgaa ctggcaggaa atgcatcaaa agacttaaag gtaaagcgta ttaccctcct 360
tcacttgcaa cttgctattc gtggagatga agaattggat tctctcatca aggtacaat 420
tgctgggtgg ggtgtcattc cacacatcca caaatctctg attgggaaga aaggacaaca 480
gaagactgtc taaaggatgc ctggattcct tgttatctca ggactctaaa tactctaaca 540
gctgtccagt gttggtgatt ccagtggact gtatctctgt gaaaaacaca attttgcctt 600
tttgtaattc tatttgagca agttggaaat ttaattagct ttccaacca ccaaatttct 660

```

```

gcattcgagt ctttaaccata ttttaagtgtt actgtggcctt caaagaagct attgattctg 720
aagtagtggg ttttgattga gttgactgtt ttttaaaaaac tgtttggatt ttaattgtga 780
tgcagaagtt atagtaacaa acatttgggtt ttgtacagac attatttcca ctctgggtga 840
taagttcaat aaaggtcata tcccaacta aaa 873

```

```

<210> 92
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002106

```

```

<400> 92
cgagtcttaa ccatatttaa gtgttactgt ggcttcaaag aagctattga ttctgaagta 60

```

```

<210> 93
<211> 4204
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002205

```

```

<400> 93
caggacaggg aagagcgggc gctatgggrga gccggacgcc agagtccctt ctccacgccc 60
tgcagctgcg ctggggccccc cggcgccgrac ccccgctcgt gccgctgctg ttgtgctcgt 120
tgccgccgcc acccaggggtc gggggcttca acttagacgc ggaggcccca gcagtactct 180
cggggccccc gggctccttc ttoggattct cagtggagtt ttaccggccg ggaacagacg 240
gggtcagtggt gctgggtggga gcacccaagg ctaataccag ccagccagga gtgctgcagg 300
gtggtgctgt ctacctctgt ccttgggggtg ccagccccac acagtgcacc cccattgaat 360
ttgacagcaa aggtctctcg ctcttggaat cctcactgtc cagctcagag ggagaggagc 420
ctgtggagta caagtccttg cagtgggtcg gggcaacagt tcgagcccat ggctcctcca 480
tcttggcatg cgctccactg tacagctgrc gcacagagaa ggagccactg agcgaccccg 540
tgggcaacctg ctacctctcc acagataa.ct tcaaccgaat tctggagtat gcacctgccc 600
gctcagatctt cagctgggca gcaggacaag gttactgcca aggaggcttc agtgccgagt 660
tcaccaagac tggccgtgtg gttttagggtg gaccaggaag ctatttctgg caaggccaga 720
tcctgtctgc cactcaggag cagattgcag aatcttatta ccccgagtac ctgatcaacc 780
tggttcaggg gcagctgcag actcgccaag ccagttccat ctatgatgac agctaccctag 840
gatactctgt ggctgttgggt gaattcagtg gtgatgacac agaagacttt gttgctggtg 900
tgcccaaagg gaacctcact tacggcta.tg tcaccatcct taatggctca gacattcgat 960
ccctctacaa cttctcaggg gaacagatgg cctcctactt tggctatgca gtggccgcca 1020
cagacgtcaa tggggacggg ctggatga.ct tgtgtgtggg ggcacccctg ctcatggatc 1080
ggacccctga cgggcggcct caggagggtgg gcagggtcta cgtctacctg cagcaccag 1140
ccggcataga gccacgccc acccttaacc tcactggcca tgatgagttt ggccgatttg 1200
gcagctcctt gacccccctg ggggacctgg accaggatgg ctacaatgat gtggccatcg 1260
gggtccctt tgggtggggag acccagca.gg gagtagtgtt tgtatttctt gggggcccg 1320
gagggtctgg ctctaagcct tcccagggtc tgcagcccct gtgggcagcc agccacaccc 1380
cagacttctt tggctctgcc cttcgaggag gccgagacct ggatggcaat ggatatcctg 1440
atctgattgt ggggtccttt ggtgtgga.ca aggtgtgtgt atacaggggc cgccccatcg 1500
tgtccgctag tgccctccctc accatcttcc ccgccatgtt caaccagag gagcggagct 1560
gcagcttaga ggggaaccct gtggcctgrca tcaaccttag cttctgcctc aatgcttctg 1620
gaaaacacgt tgctgactcc attggtttca cagtggaaact tcagctggac tggcagaagc 1680
agaagggagg ggtacggcgg gcactgttcc tggcctccag gcaggcaacc ctgaccgaga 1740
ccctgtctcat ccagaatggg gctcgaga.gg attgcagaga gatgaagatc tacctcagga 1800
acgagtcaga atttcgagac aaactctcgc cgattcacat cgctctcaac ttctccttgg 1860
acccccaaag cccagtggac agccacggcc tcaggccagc cctacattat cagagcaaga 1920
gccggataga ggacaaggct cagatcttgc tggactgtgg agaagacaac atctgtgtgc 1980
ctgacctgca gctggaagtg tttgggga.gc agaaccatgt gtacctgggt gacaagaatg 2040
ccctgaacct cactttccat gccagaa.tg tgggtgaggg tggcgccctat gaggctgagc 2100
ttcgggtcac cgcccctcca gaggctga.gt actcaggact cgtcagacac ccagggaact 2160

```

tctccagcct	gagctgtgac	tactttgccg	tgaaccagag	ccgcctgctg	gtgtgtgacc	2220
tgggcaaccc	catgaaggca	ggagccagt c	tgtgggggtgg	ccttcgggttt	acagtccctc	2280
atctccggga	cactaagaaa	accatccagt	ttgacttcca	gatcctcagc	aagaatctca	2340
acaactcgca	aagcgacgtg	gtttcctttc	ggctctccgt	ggagggtcag	gcccagggtca	2400
ccctgaacgg	tgtctccaag	cctgaggcag	tgctattccc	agtaagcgac	tggcatcccc	2460
gagaccagcc	tcagaaggag	gaggacctgg	gacctgctgt	ccaccatgtc	tatgagctca	2520
tcaaccaagg	ccccagctcc	attagccagg	gtgtgctgga	actcagctgt	ccccaggctc	2580
tggaagggtca	gcagctccta	tatgtgacca	gagttacggg	actcaactgc	accaccaatc	2640
acccatttaa	cccaaagggc	ctggagttgg	atcccagggg	ttccctgcac	caccagcaaa	2700
aacgggaagc	tccaagccgc	agctctgct t	cctcgggacc	tcagatcctg	aaatgcccgg	2760
aggctgagtg	tttcaggctg	cgctgtgagc	tcggggccct	gcaccaacaa	gagagccaaa	2820
gtctgcagtt	gcatttcoga	gtctggggcca	agactttctt	gcagcgggag	caccagccat	2880
ttagcctgca	gtgtgaggct	gtgtacaaa g	ccctgaagat	gccctaccga	atcctgcctc	2940
ggcagctgcc	caaaaaagag	cgtcagggtg	ccacagctgt	gcaatggacc	aaggcagaag	3000
gcagctatgg	cgtcccactg	tggatcatca	tcctagccat	cctgtttggc	ctcctgctcc	3060
taggtctact	catctacatc	ctctacaagc	ttggattctt	caaacgctcc	ctcccatatg	3120
gcaccgccat	ggaaaaagct	cagctcaagc	ctccagccac	ctctgatgcc	tgagtccctc	3180
caatttcaga	ctcccattcc	tgaagaacca	gtccccccac	cctcattcta	ctgaaaagga	3240
ggggctctggg	tacttcttga	aggtgctga c	ggccaggggag	aagctcctct	ccccagccca	3300
gagacatact	tgaagggcca	gagccagggg	ggtgaggagc	tggggatccc	tcccccccat	3360
gcactgtgaa	ggacccttgt	ttacacata c	cctcttcatg	gatgggggaa	ctcagatcca	3420
gggacagagg	cccagcctcc	ctgaagcct t	tgcatTTTTg	agagtttcct	gaaacaactg	3480
gaaagataac	taggaaatcc	attcacagt t	ctttggggcca	gacatgccac	aaggacttcc	3540
tgtccagctc	caacctgcaa	agatctgtcc	tcagccttgc	cagagatcca	aaagaagccc	3600
ccagtaagaa	cctggaactt	ggggagttaa	gacctggcag	ctctggacag	ccccacctg	3660
gtggggccaac	aaagaacact	aactatgcat	ggtgccccag	gaccagctca	ggacagatgc	3720
cacaaggata	gatgctggcc	cagggccaga	gccagctcc	aaggggaatc	agaactcaaa	3780
tggggccaga	tccagcctgg	ggtctggagt	tgatctggaa	cccagactca	gacattggca	3840

ccaatccagg	cagatccagg	actatatttg	ggcctgctcc	agacctgatc	ctggaggccc	3900
agttcaccct	gatttaggag	aagccaggaa	tttcccaggga	cctgaagggg	ccatgatggc	3960
aacagatctg	gaacctcagc	ctggccagac	acaggccctc	cctgttcccc	agagaaaggg	4020
gagcccactg	tcctgggcct	gcagaatttg	ggttctgcct	gccagctgca	ctgatgctgc	4080
ccctcatctc	tctgccaac	ccttcctcca	ccttggcacc	agacaccag	gacttattta	4140
aactctgttg	caagtgcaat	aaatctgacc	cagtgcctcc	actgaccaga	actagaaaaa	4200
aaaa	4204					

<210> 94
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002205

<400> 94
 ttggcaccag acaccagga cttatttaaa ctctgttgca agtgcaataa atctgaccca 60

<210> 95
 <211> 1976
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002266

<400> 95
 gccacacggt ctttgagctg agtcgagggtg gaccctttga acgcagtcgc cctacagccg 60
 ctgattcccc ccgcctcgcc tcccgtggaa gccaggccc gcttcgcagc tttctccctt 120
 tgtctcataa ccatgtccac caacgagaat gctaatacac cagctgcccg tcttcacaga 180
 ttcaagaaca agggaaaaga cagtacagaa atgaggcgctc gcagaataga ggtcaatgtg 240

```

gagctgagga aagctaagaa ggatgacCag atgctgaaga ggagaaatgt aagctcattt 300
cctgatgatg ctacttctcc gctgcaggaa aaccgcaaca accagggcac tgtaaattgg 360
tctgttgatg acattgtcaa aggcataaat agcagcaatg tggaaaatca gctccaagct 420
actcaagctg ccaggaaact actttccaga gaaaaacagc ccccatatga caacataatc 480
cgggctggtt tgattccgaa atttgtgtcc ttcttgggca gaactgattg tagtcccatt 540
cagtttgaat ctgcttgggc actcaactaac attgcttctg ggacatcaga acaaaccaag 600
gctgtggtag atggaggtgc catcccaGca ttcatcttc tggtggcatc tcccatgct 660
cacatcagtg aacaagctgt ctgggctcta ggaaacattg caggtgatgg ctcagtgttc 720
cgagacttgg ttattaagta cgggtgcaGtt gaccactgt tggtctcct tgcagttcct 780
gatatgtcat ctttagcatg tgggtacttta cgtaacttta cctggacact ttctaactct 840
tgccgcaaca agaactctgc acccccgata gatgctgttg agcagattct tcctacctta 900
gttcggctcc tgcacatga tgatccagaa gtgttagcag atacctgctg ggctatttcc 960
taccttactg atgggtccaa tgaacgaatt ggcattggtg tgaaaacagg agttgtgccc 102 O
caacttgtga agcttctagg agcttctgaa ttgccattg tgactcctgc cctaagagcc 108 O
atagggaata ttgtcactgg tacagatgaa cagactcagg ttgtgattga tgcaggagca 114 O
ctcgcctgtc ttcccagcct gctcaccAAC cccaaaacta acattcagaa ggaagctacg 120 O
tggacaatgt caaacatcac agccggc cgc caggaccaga tacagcaagt tgtgaatcat 126 O
ggattagtcc cattccttgt cagtgttctc tctaaggcag attttaagac aaaaaggaa 132 O
gctgtgtggg ccgtgaccaa ctataccagt ggtggaacag ttgaacagat tgtgtacctt 138 O
gttactgtg gcataataga accgttgatg aacctcttaa ctgcaaaaga taccaagatt 144 O
attctgggta tcctggatgc catttcaaat atctttcagg ctgctgagaa actaggtgaa 150 O
actgagaaac ttagtataat gattgaaGaa tgtggaggct tagacaaaat tgaagctcta 156 O
caaaaccatg aaaatgagtc tgtgtataag gcttcgttaa gcttaattga gaagtatttc 162 O
tctgtagagg aagaggaaga tcaaaacgtt gtaccagaaa ctacctctga aggctacact 168 O
ttccaagttc aggatggggc tcctgggacc tttaactttt agatcatgta gctgagacat 174 O
aaatttggtg tgtactacgt ttggtatttt gtcttattgt ttctctacta agaactcttt 180 O
cttaaatgtg gtttgttact gtagcac ttt ttacactgaa actatacttg aacagttcca 186 O
actgtacata catactgtat gaagcttgtc ctctgactag gtttctaatt tctatgtgga 192 O
atttcctatc ttgcagcatc ctgtaaa taa acattcaagt ccacccttaa aaaaaa 1976

```

<210> 96

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002266

<400> 96

tgagtctgtg tataaggctt cgtaaGctt aattgagaag tatttctctg tagaggaaga 60

<210> 97

<211> 1145

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002346

<400> 97

```

gctccggcca gcccggtcc agagcgCgcg aggttcgggg agctccgccca ggctgctggt 60
acctgctcc gcccggcgag caggacaggc tgctttggtt tgtgacctcc aggcaggacg 120
ccatcctct ccagaatgaa gatcttcttg ccagtgtgc tggctgccct tctgggtgtg 180
gagcgagcca gctcgtgat gtgcttctcc tgcttgaacc agaagagcaa tctgtactgc 240
ctgaagccga ccatctgtc cgaccaggac aactactgcg tgactgtgtc tgctagtgc 300
ggcattggga atctcgtgac atttggccac agcctgagca agacctgttc cccggcctgc 360
cccatcccag aaggcgtaaa tgttggTgtg gcttccatgg gcatcagctg ctgccagagc 420
tttctgtgca atttcagtgc ggccgaTggc gggctgcggg caagcgtcac cctgctgggt 480
gccgggctgc tgctgagcct gctgccggcc ctgctgcggg ttggccctg accgccaga 540
ccctgtcccc cgatcccca gctcaggaaag gaaagccag ccctttctgg atccacagt 600
gtatgggagc cctgactcc tcacgtGcct gatctgtgcc cttgggtccca ggtcaggccc 660

```

```

acccccctgca cctccacctg cccca.gcccc tgcctctgcc caagtggggc agctgccctc 720
acttctgggg tggatgatgt gacct.tcctt gggggactgc ggaagggacg aggggtccct 780
ggagtcttac ggtccaacat cagac.caagt cccatggaca tgctgacagg gtccccaggg 840
agaccgtgtc agtagggatg tgtgc.ctggc tgtgtacgtg ggtgtgcagt gcacgtgaga 900
gcacgtggcg gcttctgggg gccat.gtttg gggagggagg tgtgccagca gcctggagag 960
cctcagtccc tgtagccccc tgcc.tggca cagctgcagt cacttcaagg gcagcctttg 1020
ggggttgggg tttctgccac ttcc.gggtct aggcctgcc caaatccagc cagtccctgcc 1080
ccagcccacc ccacattgg agcc.tcctg ctgctttggt gcctcaaata aatacagatg 1140
tcccc 1145

```

<210> 98

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002346

<400> 98

```

ggttccctgg agtcttacgg tccaacatca gaccaagtcc catggacatg ctgacagggg 60

```

<210> 99

<211> 1390

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002358

<400> 99

```

gggaagtgct gttggagccg ctgt.ggttgc tgtccgcgga gtggaagcgc gtgcttttgt 60
ttgtgtccct ggccatggcg ctgcagctct cccgggagca gggaatcacc ctgcgcggga 120
gcgccgaaat cgtggccgag ttct.tctcat tcggcatcaa cagcatttta tatcagcgtg 180
gcataatatc atctgaaacc ttactctgag tgcagaaata cggactcacc ttgcttgtaa 240
ctactgatct tgagctcata aaatacctaa ataatgtggt ggaacaactg aaagattggt 300
tatacaagtg ttcagttcag aaactgggtg tagttatctc aaatattgaa agtggtgagg 360
tcctggaaag atggcagttt gatat.tgagt gtgacaagac tgcaaaagat gacagtgcac 420
ccagagaaaa gtctcagaaa gctatccagg atgaaatccg ttcagtgatc agacagatca 480
cagctacggg gacatttctg ccact.gttgg aagtttcttg ttcatttgat ctgctgattt 540
atacagacaa agatttggtt gtacc.tgaaa aatgggaaga gtcgggacca cagtttatta 600
ccaattctga ggaagtcgcg cttcg.ttcac ttactactac aatccacaaa gtaaatagca 660
tggtggccta caaaattcct gtcaatgact gaggatgaca tgaggaaaat aatgtaattg 720
taattttgaa atgtggtttt cctgaaatca ggcatctat agttgatatg ttttatttca 780
ttggttaatt tttacatgga gaaa.accaa atgatactta ctgaactgtg tgtaattggt 840
cctttatttt tttggtacct attt.gactta ccatggagtt aacatcatga atttattgca 900
cattgttcaa aaggaaccag gaggt.ttttt tgtcaacatt gtgatgtata ttcctttgaa 960
gatagtaact gtagatggaa aaact.tgtgc tataaagcta gatgctttcc taaatcagat 1020
gttttggtca agtagtttga ctcagtatat gtagggagat atttaagtat aaaatacaac 1080
aaaggaagtc taaatattca gaat.ttttgt taaggtcctg aaagtaactc ataactata 1140
aacaatgaaa tattgctgta tagct.ccttt tgaccttcat ttcattgata gttttcccta 1200
ttgaatcagt ttccaattat ttgac.tttta tttatgtaac ttgaacctat gaagcaatgg 1260
atatattgtac tctttaatgt tctgt.gatac agaactctta aaaatgtttt ttcattgtgt 1320
ttataaaatc aagttttaag tgaa.agtgag gaaataaagt taagtttggt ttaaaaaaaa 1380
aaaaaaaaa 1390

```

<210> 100

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002358

<400> 100

atgcttttccct aaatcagatg ttttgggtcaa gtagtttgac tcagtatagg tagggagata 60

<210> 101

<211> 1821

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 101

```

acaaggaggc aggcaagaca gcaaggcata gagacaacat agagctaagt aaagccagtg 60
gaaatgaaga gtcttccaat cctactgttg ctgtgcgtgg cagtttgctc agcctatcca 120
ttggatggag ctgcaagggg tgagga cacc agcatgaacc ttgttcagaa atatctagaa 180
aactactacg acctcaaaaa agatgt gaaa cagtttggtt ggagaaagga cagtggtcct 240
gttggttaaaa aaatccgaga aatgca gaag ttccctggat tggaggtgac ggggaagctg 300
gactccgaca ctctggagggt gatgcg caag ccaggtgtg gagttcctga tgttggtcac 360
ttcagaacct ttcttggcat cccgaa gtgg agggaaaacc accttacata caggattgtg 420
aattatacac cagatttgcc aaaaga tgct gttgattctg ctgttgagaa agctctgaaa 480
gtctgggaag aggtgactcc actcac attc tccaggctgt atgaaggaga ggctgatata 540
atgatctctt ttgcagttag agaaca tgga gacttttacc cttttgatgg acctggaaat 600
gttttggccc atgcctatgc cctggggcca gggattaatg gagatgcca ctttgatgat 660
gatgaacaat ggacaaagga tacaac aggg accaatttat ttctcgttgc tgctcatgaa 720
attggccact ccttgggtct ctttca.ctca gccaacactg aagctttgat gtaccactc 780
tatcactcac tcacagacct gactcgttcc cgcctgtctc aagatgatat aaatggcatt 840
cagtccctct atggacctcc cctga.ctcc cctgagacct cctgggtacc cacggaacct 900
gtccctccag aacctgggac gccagc caac tgtgatcctg ctttgtcctt tgatgctgtc 960
agcactctga ggggagaaat cctgat cttt aaagacaggc acttttggcg caaatccctc 1020
aggaagcctg aacctgaatt gcattt gatc tcttcatttt ggccatctct tccttcaggc 1080
gtggatgccg catatgaagt tactag caag gacctcgttt tcatttttaa aggaaatcaa 1140
ttctggggcca tcagaggaaa tgaggt acga gctggatacc caagaggcat ccacacccta 1200
ggtttccctc caaccgtgag gaaaat cgat gcagccattt ctgataagga aaagaacaaa 1260
acataatttct ttgtagagga caaata.ctgg agatttgatg agaagagaaa ttccatggag 1320
ccaggctttc ccaagcaaat agctga.agac ttccaggga ttgactcaaa gattgatgct 1380
gtttttgaag aatttgggtt ctttta.tttc ttactggat cttcacagt ggagtttgac 1440
ccaaatgcaa agaaagtgc acacac.tttg aagagtaaca gctggcttaa ttgttgaaag 1500
agatatgtag aaggcacaat atgggc.actt taaatgaagc taataattct tcacctaat 1560
ctctgtgaat tgaaatgttc gttttc.ctcc gcctgtgctg tgactcagat cacactcaag 1620
ggaacttgag cgtgaatctg tatctt.gccg gtcattttta tgttattaca gggcattcaa 1680

```

```

atgggctgct gcttagcttg cacctt.gtca catagagtga tctttcccaa gagaagggga 1740
agcactcgtg tgcaacagac aagtga.ctgt atctgtgtag actatttgct tatttaataa 1800
agacgatttg tcagttgttt t 1821

```

<210> 102

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 102

tgtagaaggc acaatatggg cacttt.aaat gaagctaata attcttcacc taagtctctg 60

<210> 103

<211> 2787

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 103

```

agagcggagg ccgcactcca gcactgcgca gggaccgcct tggaccgcag ttgccggcca 60
ggaatcccag tgtcacggtg gacacgcctc cctcgcgccc ttgccgcca cctgctcacc 120
cagctcaggg gctttggaat tctgtggcca cactgcgagg agatcggttc tgggtcggag 180
gctacaggaa gactcccact ccctgaaatc tggagtgaag aacgccgcca tccagccacc 240
attccaagga ggtgcaggag aacagctctg tgataccatt taacttggtg acattacttt 300
tatttgaagg aacgtatatt agagcttact ttgcaaagaa ggaagatggt tgtttccgaa 360
gtggacatcg caaaagctga tccagctgct gcatcccacc ctctattact gaatggagat 420
gctactgtgg ccagaaaaaa tccaggctcg gtggctgaga acaacctgtg cagccagtat 480
gaggagaagg tgcgcccctg catcgacctc attgaactcc tgcgggctct aggtgtggag 540
caggacctgg ccctgccagc catcgccgtc atcggggacc agagctcggg caagagctcc 600
gtgttgagg cactgtcagg agttgccctt cccagaggca gcgggatcgt gaccagatgc 660
ccgtgggtgc tgaaactgaa gaaacttgtg aacgaagata agtggagagg caaggcagat 720
taccaggact acgagattga gatttcgcat gcttcagagg tagaaaagga aattaataaa 780
gcccagaatg ccatcgccgg ggaaggaatg ggaatcagtc atgagctaat caccctggag 840
atcagctccc gagatgtccc ggatctgact ctaatagacc ttcttggtat aaccagagt 900
gctgtgggca atcagctgct tgacattggg tataagatca agacactcat caagaagtac 960
atccagaggc aggagacaat cagcctggtg gtggtcccca gtaatgtgga catcgccacc 1020
acagaggctc tcagcatggc ccaggaggtg gaccccgagg gagacaggac catcggaatc 1080
ttgacgaagc ctgatctggt ggacaaagga actgaagaca aggttgtgga cgtggtgcgg 1140
aacctcgtgt tccacctgaa gaagggttac atgattgtca agtgccgggg ccagcaggag 1200
atccaggacc agctgagcct gtccgaaacc ctgcagagag agaagatctt ctttgagaac 1260
cacccatatt tcagggatct gctggaggaa ggaaaggcca cggttccctg cctggcagaa 1320
aaacttacca gcgagctcat cacacatata tgtaaatact tgccctgtt agaaaatcaa 1380
atcaaggaga ctaccagag aataacagag gagctacaaa agtatggtgt cgacataccg 1440
gaagacgaaa atgaaaaaat gttcttctct atagataaaa ttaatgcctt taatcaggac 1500
atcactgctc tcatgcaagg agaggaaact gtaggggagg aagacattcg gctgtttacc 1560
agactccgac acgagttcca caaatggagt acaataattg aaaacaattt tcaagaaggc 1620
cataaaaatt tgagtagaaa aatccagaaa tttgaaaatc agtatcgtgg tagagagctg 1680
ccaggctttg tgaattacag gacatttgag acaactgtga aacagcaaat caaggcactg 1740
gaagagccgg ctgtggatat gctacacacc gtgacggata tggtcgggtc tgctttcaca 1800
gatgtttcga taaaaaattt tgaagagttt tttaacctcc acagaaccgc caagtccaaa 1860
attgaagaca ttagagcaga acaagagaga gaaggtgaga agctgatccg cctccacttc 1920
cagatggaac agattgtcta ctgccaggac caggtataca ggggtgcatt gcagaaggctc 1980
agagagaagg agctggaaga agaaaagaa aagaaatcct gggattttgg ggctttccag 2040
tccagctcgg caacagactc ttccatggag gagatctttc agcacctgat ggcctatcac 2100
caggaggcca gcaagcgcac ctccagccac atccctttga tcatccagtt cttcatgctc 2160
cagacgtacg gccagcagct tcagaaaggcc atgctgcagc tcctgcagga caaggacacc 2220
tacagctggc tcctgaagga gcggagcgac accagcgaca agcggaggtt cctgaaggag 2280
cggcttgca cggctgacga ggctcggcgc cggcttgccc agttccccgg ttaaccacac 2340
tctgtccagc ccgtagacg tgcacgcaca ctgtctgccc ccgttcccggt gtagccactg 2400
gactgaacgac ttgagtgtct agtagtcaga ctggatagtc cgtctctgct tatccgttag 2460
ccgtgggtgat ttagcaggaa gctgtgagag cagtttggtt tctagcatga agacagagcc 2520
ccacctcag atgcacatga gctggcggga ttgaaggatg ctgtcttcgt actgggaaaag 2580
ggattttcag cctcagaat cgctccacct tgcagctctc cccttctctg tattcctaga 2640
aactgacaca tgctgaacat cacagcttat ttctcattt ttataatgtc ccttcacaaa 2700
cccagtgttt taggagcatg agtgccgtgt gtgtgcgtcc tgtcggagcc ctgtctctc 2760
tctctgtaat aaactcattt ctacagc 2787

```

<210> 104

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 104

actgacacat gctgaacatc acagcttatt tcttcatttt tataatgtcc cttcacaaac 60

<210> 105

<211> 2808

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002759

<400> 105

gggcgccggcgg cggcgccagtt tgctcactact ttgtgacttg cggtcacagt ggcattcagc 60
 tccacactttg gtgaaccac aggcacgaca agcatagaaa catcctaaac aatcttcac 120
 gaggcatcga ggtccatccc aataaaaatc aggagaccct ggctatcata gaccttagtc 180
 ttcgctggta tactcgctgt ctgtcaacca gcgggtgact ttttttaagc cttctttttt 240
 ctctttttacc agtttctgga gcaaat tcag tttgccttc tggatttgta aattgtaatg 300
 acctcaaaac tttagcagtt cttccatctg actcagggtt gcttctctgg cggctctcag 360
 aatcaacatc cacacttccg tgattatctg cgtgcatttt ggacaaagct tccaaccagg 420
 atacgggaag aagaaatggc tgggtgatctt tcagcagggt tcttcattga ggaacttaat 480
 acataccgtc agaagcagg agtagtactt aaatatcaag aactgcctaa ttcaggacct 540
 ccacatgata ggaggttttac atttcaagtt ataatagatg gaagagaatt tccagaagggt 600
 gaaggtagat caaagaagga agcaaaaaat gccgcagcca aattagctgt tgagatactt 660
 aataaggaaa agaaggcagt tagtctctta ttattgacaa caacgaattc ttcagaagga 720
 ttatccattg ggaattacat aggccttctc aatagaattg ccagaagaa aagactaact 780
 gtaaatattg aacagtgtgc atcgggggtg catggggccag aaggatttca ttataaatgc 840
 aaaatgggag agaaagaata tagtatgggt acagggttcta ctaaacagga agcaaaaaca 900
 ttggccgcta aacttgcata tcttcagata ttatcagaag aaacctcagt gaaatctgac 960
 tacctgtcct ctggttcttt tgctactacg tgtgagtccc aaagcaactc tttagtgcac 1020
 agcacactcg cttctgaatc atcatctgaa ggtgacttct cagcagatac atcagagata 1080
 aattctaaca gtgacagttt aaacagttct tcgttgctta tgaatggtct cagaaataat 1140
 caaagggaag caaaaagatc tttggcacc agatttgacc ttcttgacat gaaagaaaca 1200
 aagtatactg tggacaagag gtttggcatg gattttaaag aaatagaatt aattggctca 1260
 ggtggatttg gccaaagttt caaagcaaaa cacagaattg acggaagac ttacgttatt 1320
 aaacgtgtta aatataataa cgagaaggcg gagcgtgaag taaaagcatt ggcaaaactt 1380
 gatcatgtaa atattgttca ctacaatggc tgttgggatg gatttgatta tgatcctgag 1440
 accagtgatg attctcttga gagcagtgat tatgatcctg agaacagcaa aaatagttca 1500
 aggtcaaaga ctaagtgcct tttcatccaa atggaattct gtgataaagg gaccttgga 1560
 caatggattg aaaaaagaag aggcgagaaa ctagacaaag ttttggcttt ggaactcttt 1620
 gaacaaataa caaaaggggt ggattatata cattcaaaaa aattaattca tagagatctt 1680
 aagccaagta atatattctt agtagataca aaacaagtaa agattggaga ctttggactt 1740
 gtaacatctc tgaaaaatga tggaaagcga acaaggagta aggggaactt gcgatacatg 1800
 agcccagaac agatttcttc gcaagactat ggaaaggag tggacctcta cgctttgggg 1860
 ctaattcttg ctgaacttct tcatgtatgt gacactgctt ttgaaacatc aaagtttttc 1920
 acagacctac gggatggcat catctcagat atatttgata aaaaagaaaa aactcttcta 1980
 cagaaattac tctcaaagaa acctgaggat cgacctaaac catctgaaat actaaggacc 2040
 ttgactgtgt ggaagaaaag ccagagaaa aatgaacgac acacatgtta gagcccttct 2100
 gaaaaagtat cctgcttctg atatgcagtt ttctttaa atctctaaat ctgctaggga 2160
 atatcaatag atatttacct tttattttta tgtttccttt aattttttac tatttttact 2220
 aatctttctg cagaaacaga aagggtttct tctttttgct tcaaaaacat tcttacattt 2280
 tactttttcc tggctcatct ctttatctct tttttttttt ttaaagacag agtctcgctc 2340
 tgttggccag gctggagtgc aatgacacag tcttggctca ctgcaacttc tgacctttgg 2400
 gttcaagtga ttctcctgcc tcagcctcct gagtagctgg attacaggca tgtgccacc 2460
 acccaactaa tttttgtgtt tttaaataaag acagggtttc accatgttgg ccaggctgggt 2520
 ctcaaactcc tgacctcaag taatccacct gcctcgccct cccaaagtgc tgggattaca 2580
 gggatgagcc accgcgcca gctcatctc tttgttctaa agatggaaaa accaccacca 2640
 aattttcttt ttatactatt aatgaatcaa tcaattcata tctatttatt aaatttctac 2700
 cgcttttagg ccaaaaaaat gtaagatcgt tctctgcctc acatagctta caagccagct 2760
 ggagaaatat ggtactcatt aaaaaaaaaa aaaaagtgtat gtacaacc 2808

<210> 106
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002759

<400> 106
 tcgttctctg cctcacatag cttacaagcc agctggagaa atatggtact cattaataaaa 60

<210> 107
 <211> 1678
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002811

<400> 107
 aagaaggagg ccgcgcgagg gctgacgaac cggaagaaga ggaactgggc ctgaaagggt 60
 accggtgacc gctactgctg ccggtgtttg cgtgtggcag ggagccaggc ctggcgagcg 120
 ggggtgtgtcg cgatgccgga gctggcagtg cagaagggtg tgggccaccc cctggtgtctg 180
 ctcagtgtgg tggatcattt caaccgaatc ggcaagggtg gaaaccagaa gcgtgttgtt 240
 ggtgtgcttt tggggtcatt gcaaaagaaa gtacttgatg tatcgaacag ttttgcagt 300
 ccttttgatg aagatgacaa agacgattct gtatggtttt tagaccatga ttatttggaa 360
 aacatgtatg gaatgtttta gaaagtcaat gccagggaaa gaatagttgg ctggtacca 420
 acaggcccta aactacacaa gaatgacatt gccatcaacg aactcatgaa aagatactgt 480
 cctaattccg tattggtcat cattgatgtg aagccgaagg acctagggct gcctacagaa 540
 gcgtacattt cagtgggaaga agtccatgat gatggaactc caacctcgaa aacatttgaa 600
 cacgtgacca gtgaaattgg agcagaggaa gctgagggaag ttggagttga acacttgta 660
 cgagatatca aagacacgac ggtgggcact ctgtcccagc ggatcacaaa ccagggtcca 720
 ggtttgaagg gactgaactc caagcttctg gatatcagga gctacctgga aaaagtgcgc 780
 acaggcaagc tgcccacaa ccaccagatc atctaccagc tgcaggacgt cttcaacctg 840
 ctgccagatg tcagcctgca ggagttcgtc aaggcctttt acctgaagac caatgaccag 900
 atggtggtag tgtacttggc ctcgctgacg cgttccgtgg tcgccctgca caacctcat 960
 aacaacaaga ttgccaaccg ggatgcagag aagaaagaag ggcaggagaa agaagagagc 1020
 aaaaaggata ggaaagagga caaggagaaa gataaagata aggaaaagag tgatgtaaa 1080
 aaagaggaga aaaaggagaa aaagtaaaac atgtattaaa tagctttttt aatttgttaa 1140
 ttaaaatctt acaaactaaa tcagtgtgct gctagagggt tctttttcac ttgacatgc 1200
 tattagaaaag ctgacccaac aagagctctc tgcctccggt cactcttgtc gtggtgcta 1260
 gtggaagtga atggagactg atctcaaata tgaactgcag ctttcgctgc tgtgagttgg 1320
 ggatatgata gtcagctcag gcttcagatt gtatgagaaa aatgaagaga agtcaacaaa 1380
 tatttttgga ctcttcattc atttatctct aaaaccagga gttgaatttt cctcatcttg 1440
 aaagactctt ggggtctgtt tctggtattt tacaaaattg ctaagtggaa tgcataaat 1500
 gcattatgtt ctctggtaac acgtagagtt cagacccttc tgaactctgt tgataatac 1560
 acaccatgtt ctggacccat agctctggca tcctcagggg ttgtgatcca gctccatat 1620
 ttgtttacct tcaaagatac aattaaatgg cttgattttt aaaaaaaaaa aaaaaaaa 1678

<210> 108
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002811

<400> 108
 aaattgctaa gtggaatgca tgaattgcat tatgttctct ggtaaacacgt agagttcaga 60

<210> 109

<211> 846
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002888

<400> 109
 ccacgtccgg ggtgccgagc caacttttcc ggcgtccatgc agccccgccg gcaacggctg 60
 cccgctccct ggtccggggc cagggggccc cgccccaccg ccccgctgct cgcgctgctg 120
 ctgttgctcg ccccggtggc ggcgcccgcg gggtcggggg gccccgacga ccctgggcag 180
 cctcaggatg ctgggggtccc gcgcaggctc ctgcagcaga aggcgcgcgc ggcgcttcac 240
 ttcttcaact tccgggtccgg ctgcgccagc gcgctgcgag tgcctggccga ggtgcaggag 300
 ggccgcgcgt ggattaatcc aaaagaggga tgtaaagtcc acgtggtctt cagcacagag 360
 cgctacaacc cagagtcctt acttcaggaa ggtgagggac gtttggggaa atgttctgct 420
 cgagtgtttt tcaagaaatca gaaacccaga ccaaccatca atgtaacttg tacacggctc 480
 atcgagaaaa agaaaagaca acaagaggat tacctgcttt acaagcaaat gaagcaactg 540
 aaaaaccctc tggaaatagt cagcatacct gataatcatg gacatattga tccctctctg 600
 agactcatct gggatttggc ttctcttgga agctcttacg tgatgtggga aatgacaaca 660
 caggtgtcac actactactt ggcacagctc actagtgtga ggcagtgggt aagaaaaacc 720
 tgaaaattaa cttgtgcacac aagagttaca atcaaagtgg tctccttaga ctgaattcat 780
 gtgaacttct aatttcatat caagagttgt aatcacattt atttcaataa atatgtgagt 840
 tcctgc 846

<210> 110
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_002888

<400> 110
 aaagaaaaga caacaagagg attacctgct ttacaagcaa atgaagcaac tgaaaaaccc 60

<210> 111
 <211> 1054
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003090

<400> 111
 gaattccgcg ggaggccacg ggctttccac agcgcggggg aacgggaggc tgcaggatgg 60
 tcaagctgac ggcggagctg atcgagcagg cggcgcagta caccaacgcg gtgcgcgacc 120
 gggagctgga cctccggggg tataaaaattc ccgtcattga aaatctaggt gctacgttag 180
 accagtttga tgcatttgat ttttctgaca atgagatcag gaaactggat ggttttccct 240
 tggtgagaag actgaaaaca ttgttagtga acaacaacag aatatgccgt ataggtgagg 300
 gacttgatca ggcctctgcc tgtctgacag aactcattct caccaataat agtctcgtgg 360
 aactgggtga tctggaccct ctggcatctc tcaaatcgct gacttaccta agtatcctaa 420
 gaaatccggt aaccaataag aagcattaca gattgtatgt gatttataaa gttccgcaag 480
 tcagagtaact ggatttcag aaagtgaac taaaagagcg tcaggaagca gagaaaatgt 540
 tcaagggcaa acgggtgca cagcttgcaa aggatattgc caggagaagc aaaactttta 600
 atccagggtc tggtttgcca actgacaaaa agagaggtgg gccatctcca ggggatgtag 660
 aagcaatcaa gaatgccata gcaaatgctt caactctggc tgaagtggag aggctgaagg 720
 ggttgctgca gtctggtcag atccctggca gagaacgcag atcagggccc actgatgatg 780
 gtgaagaaga gatggaagaa gacacagtca caaacgggtc ctgagcagtg aggcagatgt 840
 ataataatag gccctcttgg aacaagtctt gcttttcgaa catggtataa tagccttgtt 900

tgtgttagca aagtgggaatc tatcagcatt gttgaaatgc ttaagactgc tgctgataat 960
 tttgtaatat aagttttgaa atctaaatgt caatttttcta caaattataa aaataaactc 1020
 cactctctat gctaaaaaaa aaaaaaagga attc 1054

<210> 112
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003090

<400> 112
 taatagcctt gtttgtgtta gcaaagtgga atctatcagc attgttgaaa tgcttaagac 60

<210> 113
 <211> 2033
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003158

<400> 113
 gaattccggg actgagctct tgaagacttg ggtccttggg cgcagggtgga ggcacgggtc 60
 tcactccatt gcccaggcca gagtgcggga tatttgataa gaaacttcag tgaaggccgg 120
 gcgcgggtgct catgcccgtat atcccagcat ttccggaggc cgaggcatca tggaccgac 180
 taaagaaaaac tgcattttcag gacctgttaa ggctacagct ccagttggag gtccaaaacg 240
 tgtttctcgtg actcagcaat ttcttctgtca gaatccatta cctgtaaata gtggccaggc 300
 tcagcgggtc ttgtgtcctt caaattcttc ccagcgcgtt cttttgcaag cacaaaagct 360
 tgtctccagt cacaagccgg ttcagaatca gaagcagaag caattgcagg caaccagtgt 420
 acctcatcct gtctccaggc cactgaataa cacccaaaaag agcaagcagc ccctgccatc 480
 gcacctgaaa ataactcctga ggaggaaactg gcatcaaaaac agaaaaatga agaatacaaaa 540
 agaggcagtg gctttggaag actttgaaat tgggtcgcct ctgggttaaag gaaagtgttg 600
 taatgtttat ttggcaagag aaaagcaaaag caagtttatt ctggctctta aagtgttatt 660
 taaagctcag ctggagaaaag ccggagtgga gcatcagctc agaagagaag tagaaataca 720
 gtccacacctt cggcatccta atattcttag actgtatggg tatttccatg atgctaccag 780
 agtctaccta attctggaat atgcaccact tgggaacagtt tatagagaac ttcagaaact 840
 ttcaaagttt gatgagcaga gaactgctaa cttatataac agaattgcaa atgccctgtc 900
 ttactgtcat tcgaagagag ttattcatag agacattaag ccagagaact tacttcttgg 960
 atcagctgga gagcttaaaa ttgcagattt tgggtggtca gtacatgctc catcttccag 1020
 gaggaccact ctctgtggca ccttggaacta cctgccccct gaaatgattg aaggtcggat 1080
 gcatgatgag aaggtggatc tctggagcct tggagttctt tgctatgaat ttttagttgg 1140
 gaagcctcct tttgaggcaa acacatacca agagacctac aaaagaatat cacgggttga 1200
 attcacattc cctgactttg taacagaggg agccaggggac ctcathttcaa gactgttgaa 1260
 gcataatccc agccagaggc caatgctcag agaagtactt gaacacccct ggatcacagc 1320
 aaattcatca aaaccatcaa attgccaaaa caaagaatca gctagcaaac agtcttagga 1380
 atcgtgcagg gggagaaatc cttgagccag ggctgccata taacctgaca ggaacatgct 1440
 actgaagttt attttaccat tgactgctgc cctcaatcta gaacgctaca caagaaatat 1500
 tttgttttta ctacagcagg gtgccttaac ctccctattc agaaagctcc acatcaataa 1560
 acatgacact ctgaagtga agtagccacg agaattgtgc tacttatact ggaacataat 1620
 ctggaggcaa ggctcgactg cagtcgaacc ttgcctccag attatgaacc agtataagta 1680
 gcacaattct cgtggctact ttcacttcag agtgtcatgt ttattgatgt ggagctttct 1740
 gaataggagg gtttaaggcac acctgctgag taaaacaaat atttcttgtg tagcgttctt 1800
 aggaatctgg tgtctgtccg gcccgggtag gcctgttggg tttctagtcc tccttaccat 1860
 catctccata tgagagtgtg aaaataggaa cacgtgctct acctccattt agggatttgc 1920
 ttgggataca gaagaggcca tgtgtctcag agctgttaag ggcttatttt tttaaaacat 1980
 tggagtcata gcatgtgtgt aaactttaaa tatgcaggcc ttctgtggctc gag 2033

<210> 114
 <211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_003158

<400> 114
ttgggttttct agtcctcctt accatcatct ccatatgaga gtgtgaaaat aggaacacgt 60

<210> 115
<211> 1421
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003258

<400> 115
acttactgcg ggacggcctt ggagagtact cgggttcgtg aacttcccgg aggcgcaatg 60
agctgcatta acctgcccac tgtgtctgcc ggctcccca gcaagaccg ggggcagatc 120
caggtgattc tcgggccgat gttctcagga aaaagcacag agttgatgag acgcgtccgt 180
cgcttccaga ttgtctagta caagtgcctg gtgatcaagt atgccaaaga cactcgctac 240
agcagcagct tctgcacaca tgaccggaac accatggagg cgctgcccgc ctgcctgctc 300
cgagacgtgg cccaggaggc cctgggcgtg gctgtcatag gcatcgacga ggggcagttt 360
ttccctgaca tcatggagtt ctgcgaggcc atggccaacg ccgggaagac cgtaattgtg 420
gctgcactgg atgggacctt ccagaggaag ccatttgggg ccattcctgaa cctgggtgccg 480
ctggccgaga cgtgggtgaa gctgacggcg gtgtgcatgg agtgcttccg ggaagccgcc 540
tataccaaga ggctcggcac agagaaggag gtcgaggtga ttggggggagc agacaagtaac 600
cactccgtgt gtcggctctg ctacttcaag aaggcctcag gccagcctgc cgggccggac 660
aacaagaga actgcccagt gccaggaaag ccagggggaag ccgtggctgc caggaagctc 720
tttgcccac agcagattct gcaatgcagc cctgccaaact gagggacctg caagggccgc 780
ccgtccctt cctgccactg ccgcctactg gacgtgcccc tgcattgctgc ccagccactc 840
caggaggaag tcgggaggcg tggagggtga ccacacctg gccttctggg aactctcctt 900
tgtgtggctg cccacactgc cgcattctcc ctctctcct accactgggt ctgcttaaaag 960
cttccctctc agctgctggg acgatcgccc aggcctggagc tggccccgct tgggtggcctg 1020
ggatctggca cactccctct ccttggggtg agggacagag cccacgctg ttgacatcag 1080
cctgcttctt cccctctgcg gcttttactg ctgagtttct gttctccctg ggaagcctgt 1140
gccagcacct ttgagccttg gccacactg aggccttaggc ctctctgcct gggatgggct 1200
cccaccctcc cctgaggatg gcctggattc acgcctctct gtttcctttt gggctcaaag 1260
cccttcctac ctctgggtgat ggtttccaca ggaacaacag catctttcac caagatgggt 1320
ggcaccaacc ttgctgggac ttggatccca ggggcttatc tcttcaagtg tggagagggc 1380
agggtccacg cctctgctgt agcttatgaa attaactaat t 1421

<210> 116
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003258

<400> 116
cttcctacct ctgggtgatgg tttccacagg aacaacagca tctttcacca agatggggtg 60

<210> 117
<211> 913
<212> DNA
<213> Homo sapiens

<300>
<308> NM_003311

<400> 117

```

agagccggcg cgcacccgc cgcattgcc gctcccagtc ccgcgctcgg caccaca tga 60
aatccccga cgaggtgcta cgcgagggcg agttggagaa gcgcagcgac agcctct tcc 120
agctatggaa gaagaagcgc ggggtgctca cctccgaccg cctgagcctg tccccgc ca 180
gccccgcgc gcgccccaa gaggctgcgt tccactccat cctcaaggtg gactgcgtgg 240
agcgacggg caagtacgtg tacttcacca tcgtcaccac cgaccacaag gagatcgact 300
tccgctgcgc gggcgagagc tgctggaacg cggccatcgc gctggcgctc atcgatt tcc 360
agaaccgcgc cgccctgcag gactttcgca gccgccagga acgcaccgca cccgcgcac 420
ccgcccagga cgcctggct gccgcggcg ccgcaccctc cgagccctcg gagccct cca 480
ggccatcccc gcagcccaaa cccgcgcgc catgagcccg ccgcgggcca tacgctggac 540

```

```

gagtcggacc gaggctagga cgtggccggc gctctccagc cctgcagcag aagaact tcc 600
cgtgcgcgcg gatcctcgct ccgttgacg ggcgccttaa gttattggac tatctaatat 660
ctatgtatatt atttcgctgg ttctttgtag tcacatatatt tatagtctta atatctt gtt 720
tttgcattcac tgtgcccatt gcaataaat caacttgcca gtttgctttt ctaccat ccg 780
gctgtggctc agtgagact ctgctgggag ggtggaggcc caggaatggg cgggcaggac 840
acctcatcc agtcctgcgc ggctgggtgt aaaggcgctg ggaaccggct ttgaatgaat 900
aatgaatcg tgt 913

```

<210> 118

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003311

<400> 118

```

atttcgctgg ttctttgtag tcacatatatt tatagtctta atatctt gtt tttgcattcac 60

```

<210> 119

<211> 1723

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003376

<400> 119

```

tcgcggaggc ttggggcagc cgggtagctc ggaggctcgt gcgctggggg ctagcaccag 60
cgctctgtcg ggaggcgca cgggttaggtg gaccggctcag cggactcacc ggccaggggcg 120
ctcgggtgctg gaatttgata ttcattgata cgggttttat cctctctctt tttcttcaa 180
catttttttt taaaactgta ttggttctcg ttttaattta tttttgcttg ccattcccc 240
cttgaatcgc gccgacggc tggggagatt gctctacttc cccaaatcac tgtggatttt 300
ggaaaccagc agaaagagga aagaggtagc aagagctcca gagagaagtc gaggaagaga 360
gagacggggg cagagagagc gcgcgggctg gcgagcagcg aaagcgacag gggcaaagtg 420
agtgacctgc ttttgggggt gaccgcggga gcgcggcgct agccctcccc cttgggatacc 480
cgcagctgac cagtcgcgc t gacggacaga cagacagaca ccgccccag cccagctac 540
cacctcctcc ccggccggcg gcggacagtg gacgcggcgg cgagccgcgg gcaggggccc 600
gagcccgcgc ccggaggcg ggtggagggg gtcggggctc gcggcgctgc actgaaactt 660
ttcgtccaac ttctgggctg ttctcgcttc ggaggagccg tggctcgcgc gggggaagcc 720
gagccgagcg gagccgcga agtgcttagc tcgggcccgg aggagccgca gccgaggag 780
ggggaggagg aagaagagaa ggaagaggag agggggccgc agtggcgact cggcgctcgg 840
aagccgggct catggacggg tgaggcggcg gtgtgcgcag acagtgtctc agccgcgcgc 900
gctccccagg ccttgggccc ggctcgggc cggggaggaa gaggtagctc ccgaggcgcc 960
gaggagagcg ggccgcccc aagcccgagc cggagaggga gcgcgagccg cgcgggccc 1020
ggtcgggcct ccgaaaccat gaactttctg ctgtcttggg tgcatggag ccttgccctg 1080
ctgctctacc tccaccatgc caagtggctc caggctgcac ccatggcaga aggaggagg 1140
cagaatcatc acgaagtggg gaagtctcat gatgtctatc agcgcagcta ctgcatcca 1200
atcgagaccc tgggtggaca ctccaggag taccctgatg agatcgagta catcttcaag 1260

```



```

ccatcctgtg  tgccctgat  gcgatgcggg  ggctgctgca  atgacgaggg  cctggagtgt  1320
gtgcccactg  aggagtccaa  catcaccatg  cagattatgc  ggatcaaacc  tcaccaaggc  1380
cagcacatag  gagagatgag  cttcctacag  cacaacaaat  gtgaatgcag  accaaagaaa  1440
gatagagcaa  gacaagaaaa  aaaatcagtt  cgaggaaagg  gaaaggggca  aaaacgaaa  1500
cgcaagaaat  cccggtataa  gtccctggagc  gttccctgtg  ggcccttgctc  agagcggaga  1560
aagcatttgt  ttgtacaaga  tccgcagacg  tgtaaattgt  cctgcaaaaa  cacagactcg  1620
cgttgcaagg  cgaggcagct  tgagttaaac  gaacgtactt  gcagatgtga  caagccgagg  1680
cgggtgagccg  ggcaggagga  aggagcctcc  ctcagggttt  cgg  1723

```

<210> 120

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003376

<400> 120

```

ccagcacata  ggagagatga  gcttcctaca  gcacaacaaa  tgtgaatgca  gaccaaagaa  60

```

<210> 121

<211> 2834

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003406

<400> 121

```

gccactccc  accgccagct  ggaaccctgg  ggactacgac  gtccctcaaa  ccttgcttct  60
aggagataaa  aagaacatcc  agtcatggat  aaaaatgagc  tggttcagaa  ggccaaaactg  120
gccgagcagg  ctgagcgata  tgatgacatg  gcagcctgca  tgaagtctgt  aactgagcaa  180
ggagctgaat  tatccaatga  ggagaggaat  ctctctcag  ttgcttataa  aaatgttgta  240
ggagcccgtat  ggtcatcttg  gagggtcgct  tcaagtattg  aacaaaagac  ggaagggtgct  300
gagaaaaaac  agcagatggc  tcgagaatac  agagagaaaa  ttgagacgga  gctaagagat  360
atctgcaatg  atgtactgtc  tcttttgga  aagttcttga  tccccaatgc  ttcacaagca  420
gagagcaaag  tcttctatct  gaaaatgaaa  ggagattact  accgttactt  ggctgaggtt  480
gccgctgggtg  atgacaagaa  agggattgtc  gatcagtcac  aacaagcata  ccaagaagct  540
tttgaaatca  gcaaaaagga  aatgcaacca  acacatccta  tcagactggg  tctggccctt  600
aacttctctg  tgttctatta  tgagattctg  aactccccag  agaaaagcctg  ctctcttgca  660
aagacagctt  ttgatgaagc  cattgctgaa  ctgatacat  taagtgaaga  gtcatacaaa  720
gacagcacgc  taataatgca  attactgaga  gacaacttga  cattgtggac  atcgataacc  780
caagagagacg  aagctgaagc  aggagaagga  ggggaaaatt  aaccggcctt  ccaacttttg  840
tctgctcat  tctaaaattt  acacagtaga  ccatttgtca  tccatgctgt  cccacaaata  900
gttttttgtt  tacgatttat  gacagggtta  tgttacttct  atttgaattt  ctatatttcc  960
catgtgggtt  ttatgtttaa  tattagggga  gtagagccag  ttaacattta  gggagttatc  1020
tgttttcatc  ttgaggtggc  caatatgggg  atgtggaatt  tttatacaag  ttataagtgt  1080
ttggcatagt  acttttggtt  catttgtggc  tcaaaagggc  cagtgtaaaa  ctgcttccat  1140
gtctaagcaa  agaaaactgc  ctacatactg  gtttgtcctg  gcggggaata  aaagggatca  1200
ttggttccag  tcacaggtgt  agtaattgtg  ggtactttta  ggtttggagc  acttacaagg  1260
ctgtggtaga  atcatacccc  atggatacca  catattaaac  catgtatatc  tgtggaatac  1320
tcaatgtgta  cacctttgac  tacagctgca  gaagtgttcc  tttagacaaa  gttgtgaccc  1380
attttactct  ggataagggt  agaaacggtt  cacattccat  tatttgtaaa  gttacctgct  1440
gttagctttc  attatttttg  ctacactcat  tttatttgta  tttaaatgtt  ttaggcaacc  1500
taagaacaaa  tgtaaaagta  aagatgcagg  aaaaatgaat  tgcttggtat  tcattacttc  1560
atgtatatca  agcacagcag  taaaacaaaa  acccatgtat  ttaacttttt  ttaggatatt  1620
ttgcttttgt  gatttttttt  tttttttttt  gatacttgcc  taacatgcat  gtgctgtaaa  1680
aatagttaac  agggaaataa  cttgagatga  tggctagctt  tgtttaatgt  cttatgaaat  1740
tttcatgaac  aatccaagca  taattgttaa  gaacacgtgt  attaaattca  tgtaagtggg  1800
ataaaagttt  tatgaatgga  cttttcaact  actttctcta  cagcttttca  tgtaaattag  1860

```

tcttggttct	gaaacttctc	taaaggaaat	tgtacatttt	ttgaaattta	ttccttattc	1920
cctcttggca	gctaattggc	tcttaccaag	tttaaacaca	aaatttatca	taacaaaaat	1980
actactaata	taactactgt	ttccatgtcc	catgatcccc	tctcttcctc	cccacccctga	2040
aaaaaatgag	ttcctatttt	ttctggggaga	ggggggggatt	gattagaaaa	aaatgtagtg	2100
tgttccattt	aaaattttgg	catatggcat	tttctaactt	aggaagccac	aatgttcttg	2160
gccccatcatg	acattgggta	gcattaactg	taagttttgt	gcttccaaat	cacttcttgg	2220
tttttaagaa	tttcttgata	ctcttatagc	ctgccttcaa	ttttgatcct	ttattctttc	2280
tatttgtcag	gtgcacaaga	ttaccttcct	gttttagcct	tctgtcttgt	caccaaccat	2340
tcttacttgg	tggccatgta	cttggaaaaa	ggccgcatga	tctttctggc	tccaactcagt	2400
gtctaaggca	ccctgcttcc	tttgcttgca	tcccacagac	tatttccctc	atcctattta	2460
ctgcagcaaa	tctctcctta	gttgatgaga	ctgtgtttat	ctccctttaa	aaccttacct	2520
atcctgaatg	gtctgtcatt	gtctgccttt	aaaatccttc	ctctttcttc	ctcctctatt	2580
ctctaaataa	tgatggggct	aagttatacc	caaagctcac	ttacaaaaat	atttctctcag	2640
tactttgcag	aaaacaccaa	acaaaaatgc	catttttaaaa	aaggtgtatt	ttttctttta	2700
gaatgtaagc	tctcaagrag	cagggacaat	gttttctgta	tgttctattg	tgcctagtac	2760

actgtaaatg	ctcaataaat	attgatgatg	ggaggcagtg	agtcttgatg	ataagggtga	2820
gaaactgaaa	tccc	2834				

<210> 122

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003406

<400> 122

tttagccttc	tgtcttgtca	ccaaccattc	ttacttgggtg	gccatgtact	tggaaaaagg	60
------------	------------	------------	-------------	------------	------------	----

<210> 123

<211> 1938

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003504

<400> 123

gatttggcgg	gagtcttgac	cgccgcgggg	ctcttgggtac	ctcagcgoga	gcgcCaggcg	60
tccggccgcc	gtggctatgt	tcgtgtccga	tttccgcaaa	gagttctacg	aggtgggtcca	120
gagccagagg	gtccttctct	tcgtggcctc	ggacgtggat	gctctgtgtg	cgtgCaagat	180
ccttcaggcc	ttgttccagt	gtgaccacgt	gcaatatacg	ctggttccag	tttcTgggtg	240
gcaagaactt	gaaactgcct	ttcttgagca	ttaaagaacag	tttcattatt	ttattctcat	300
aaactgtgga	gctaattgtag	acctatttga	tattcttcaa	cctgatgaag	acactatatt	360
ctttgtgtgt	gacacccata	ggccagtcac	tgtcgtcaat	gtatacaacg	atacCagat	420
caaattactc	attaaacaag	atgatgacct	tgaagttccc	gcctatgaag	acatcttcag	480
ggatgaagag	gaggatgaag	agcattcagg	aaatgacagt	gatgggtcag	agccTcttga	540
gaagcgcaca	cggttagaag	aggagatagt	ggagcaaacc	atgcccaggga	ggcagcgggc	600
agagtgggag	gcccggagaa	gagacatcct	ctttgactac	gagcagtatg	aataTcatgg	660
gacatcgtca	gccatggtga	tgtttgagct	ggcttggatg	ctgtccaagg	acctgaatga	720
catgctgtgg	tgggccatcg	ttggactaac	agaccagtgg	gtgcaagaca	agatCactca	780
aatgaaatac	gtgactgatg	ttggtgtcct	gcagcgccac	gtttcccggc	acaaCcaccc	840
gaacgaggat	gaggagaaca	cactctccgt	ggactgcaca	cggatctcct	ttgagtatga	900
cctccgcctg	gtgctctacc	agcactggtc	cctccatgac	agcctgtgca	acacCagcta	960
taccgcagcc	aggttcaagc	tgtgggtctgt	gcatggacag	aagcggctcc	aggagtctct	1020
tgcagacatg	ggtcttccc	tgaagcaggt	gaagcagaag	ttccaggcca	tggacatctc	1080
cttgaaggag	aatttgcggg	aaatgattga	agagtctgca	aataaatttg	ggatgaagga	1140
catgcgcgtg	cagactttca	gcattcattt	tgggttcaag	cacaagtttc	tggcCagcga	1200
cgtggctctt	gccaccaTgt	ctttgatgga	gagccccgag	aaggatggct	caggGacaga	1260
tcacttcac	caggctctgg	acagcctctc	caggagtaac	ctggacaagc	tgtacCatgg	1320

```

cctggaactc gccaagaagc agctgcgagc caccacagcag accattgcca gctgcctttg 1380
caccaacctc gtcactctcc aggggccttt cctgtactgc tctctcatgg agggcactcc 1440
agatgtcatg ctgttctcta ggccggcatc cctaagcctg ctcagcaaac acctgcctcaa 1500
gtcctttgtg tgttcgacaa agaaccggcg ctgcaaactg ctgccccctgg tgatggctgc 1560
ccccctgagc atggagcatg gcacagtgc cgtgggtgggc atccccccag agaccgacag 1620
ctcggacagg aagaactttt ttgggagggc gtttgagaag gcagcggaaa gcaccagctc 1680
ccggatgctg cacaaccatt ttgacctctc agtaattgag ctgaaagctg aggatcggag 1740
caagtttctg gacgcactta tttccctcct gtcctaggaa tttgattctt ccagaatgac 1800
cttcttattt atgtaactgg ctttcattta gattgtaagt tatggacatg atttgagatg 1860
tagaagccat tttttattaa ataaaatgct tatttttaggc tccgtcccca aaaaaaaaaa 1920
aaaaaaaaaa aaaaaaaa 1938

```

```

<210> 124
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003504

```

```

<400> 124
caagtttctg gacgcactta tttccctcct gtcctaggaa tttgattctt ccagaatgac 60

```

```

<210> 125
<211> 2346
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003600

```

```

<400> 125
acaaggcagc ctgcgtcgag cgcaggccaa tcggctttct agctagaggg tttaa.ctcct 60
atttaaaaag aagaaccttt gaattctaac ggctgagctc ttggaagact tgggt.ccttg 120
ggtcgcagggt gggagccgac ggggtgggtag accgtggggg atatctcagt ggccgacgag 180
gacggcgggg acaaggggag gctggtcgga gtggcggagc gtcaagtccc ctgtc.ggttc 240
ctcgtctcct gagtgtcctt ggcgtgcct tgtgcccgcc cagcgccttt gcctc.cgctc 300
ctgggcaccg aggcgccttg taggatactg cttgttactt attacagcta gaggc.atcat 360
ggaccgatct aaagaaaact gcatttcagg acctgttaag gctacagctc cagtt.ggagg 420
tccaaaacgt gttctcgtga ctcagcaatt tccttgtcag aatccattac ctgtaaataag 480
tggccaggct cagcgggtct tgtgtccttc aaattcttcc cagcgcattc ctttg.caagc 540
acaaaagctt gtctccagtc acaagccggt tcagaatcag aagcagaagc aattg.caggc 600
aaccagtgta cctcatcctg tctccaggcc actgaataac acccaaaaaga gcaag.cagcc 660
cttgccatcg gcacctgaaa ataatcctga ggaggaactg gcatcaaaac agaaa.aatga 720
agaatcaaaa aagaggcagt gggctttgga agactttgaa attggtcgcc ctctg.ggtta 780
aggaaagttt ggtaatgttt atttggcaag agaaaagcaa agcaagttaa ttctg.gctct 840
taaagtgtta tttaaagctc agctggagaa agccggagtg gagcatcagc tcaga.agaga 900
agtagaaata cagtcccacc ttcggcatcc taatattctt agactgtatg gttat.tcca 960
tgatgctacc agagtctacc taattctgga atatgcacca cttggaacag tttat.agaga 1020
acttcagaaa ctttcaaagt ttgatgagca gagaactgct acttatataa cagaa.ttggc 1080
aaatgccttg tcttactgtc attcgaagag agttattcat agagacatta agcca.gagaa 1140
cttacttctt ggcacagctg gagagcctaa aattgcagat tttgggtggg cagta.catgc 1200
tccatcttcc aggaggacca ctctctgtgg caccctggac tacctgcccc ctgaa.atgat 1260
tgaaggctcg atgcatgatg agaaggtgga tctctggagc cttggagttc tttgc.tatga 1320
attttttagt gggaagcctc cttttgaggg aaacacatac caagagacct acaaa.agaat 1380
atcacgggtt gaattcacat tccctgactt tgtaacagag ggagccaggg acctc.atttc 1440
aagactgttg aagcataatc ccagccagag gccaatgctc agaga.agtac ttgaa.caccc 1500
ctggatcaca gcaaatctat caaaaccatc aaattgccaa acaaa.agaat 1560
acagtcttag gaatcgtgca gggggagaaa tccttgagcc agggctgcca tataa.cctga 1620
caggaacatg ctactgaagt ttattttacc attgactgct gccctcaatc tagaa.cgcta 1680
cacaagaaat atttgtttta ctcagcagggt gtgccttaac ctccctattc agaa.agctcc 1740

```

```

acatcaataa acatgacact ctgaagtgaag agtagccacg agaattgtgc tacttatact 1800
ggttcataat ctggaggcaa ggctcgactg cagccgcccc gtcagcctgt gctaggcatg 1860
gtgtcttcac agrgaggcaaa tccagagcct ggctgtgggg aaagtgacca ctctgccctg 1920
accccgatca gttaaggagc tgtgcaataa ccttcctagt acctgagtga gtgtgtaact 1980
tattgggttg gcgaagcctg gtaaagctgt ttggaatgagt atgtgattct ttttaagtat 2040
gaaaataaag atatatgtac agacttgat tttttctctg gtggcattcc ttttagaatg 2100
ctgtgtgtct gtccggcacc ccggtaggcc tgattgggtt tctagtcctc cttaccact 2160
tatctcccat atgagagtgt gaaaaatagg aacacgtgct ctacctccat ttagggattt 2220
gcttgggata cagaagaggc catgtgtctc agagctgtta agggcttatt tttttaaac 2280
attggagtca tagcatgtgt gtaaacttta aatatgcaaa taaataagta tctatgtcta 2340
aaaaaa 2346

```

<210> 126

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003600

<400> 126

```

agagtgtgaa aaataggaac acgtgctcta cctccattta gggatttgct tgggatacag 60

```

<210> 127

<211> 853

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003641

<400> 127

```

ctagtctga ctctacttct gatgaggaag cctctctcct tagccttcag cctttcctcc 60
caccctgcca taagtaattt gatcctcaag aagttaaacc acacctcatt ggctcctggc 120
taattcacca atttacaac agcaggaaat agaaacttaa gagaaataca cacttctgag 180
aaactgaaac gacaggggaa aggaggtctc actgagcacc gtccagcat ccggacacca 240
cagcggccct tcgctccacg cagaaaaacca cacttctcaa accttcactc aacacttctc 300
tccccaaagc cagaagatgc acaaggagga acatgagggt gctgtgctgg gggcaccgcc 360
cagcaccatc ctccaaggt ccaccgtgat caacatccac agcgagacct ccgtgcccga 420
ccatgtcgtc tggctcctgt tcaacaccct cttcttgaac tgggtgctgtc tgggcttcat 480
agcattcgcc taactcgtga agtctaggga caggaagatg gttggcgacg tgaccggggc 540
ccaggcctat gcctccaccg ccaagtgcct gaacatctgg gccctgattc tgggcatact 600
catgaccatt ggattcatcc tgtcactggt attcggtctc gtgacagtct accatattat 660
gttacagata atacaggaaa aacgggggtta ctagtagccg cccatagcct gcaacctttg 720
cactccactg tgcaatgctg gccctgcacg ctggggctgt tggccctgcc ccttggtcc 780
tgcccctaga tacagcagtt tatacccaca cacctgtcta cagtgtcatt caataaagtg 840
cacgtgcttg tga 853

```

<210> 128

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003641

<400> 128

```

attatgttac agataatata ggaaaaacgg ggttactagt agccgcccat agcctgcaac 60

```

<210> 129
 <211> 1280
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003756

<400> 129
 gaaagatggc gtccc gcaag gaaggtaccg gctctactgc cacctcttcc agc tccaccg 60
 ccggcgcagc agggaaaggc aaaggcaaag gcggctcggg agattcagcc gtgaagcaag 120
 tgcagataga tggccttgtg gtattaaaga taatcaaaca ttatcaagaa gaaaggacaag 180
 gaactgaagt tgttcaagga gtgcttttgg gtctggttgt agaagatcgg cttgaaatta 240
 ccaactgctt tcctt tccct cagcacacag aggatgatgc tgactttgat gaaagtccaat 300
 atcagatgga aatga tgcgg agccttcgcc atgtaaacad tgatcatctt cactgtgggct 360
 ggtatcagtc cacat actat ggctcattcg ttaccggggc actcctggac tctcagttta 420
 gttaccagca tgccattgaa gaatctgtcg ttctcattta tgatcccata aaaactgccc 480

 aaggatctct ctact taaag gcatacagac tgactcctaa actgatggaa gtt tgtaaag 540
 aaaaggattt ttccc ctgaa gcattgaaaa aagcaaatat cacctttgag tacatgtttg 600
 aagaagtgcc gattgtaatt aaaaattcac atctgatcaa tgtcctaata tgggaacttg 660
 aaaagaagtc agctgttgca gataaacatg aattgctcag ccttgccagc agcaatcatt 720
 tggggaagaa tctac agttg ctgatggaca gagtggatga aatgagccaa gat atagtta 780
 aatacaacac ataca tgagg aatactagta aacaacagca gcagaaacat cagtatcagc 840
 agcgtcgcca gcaggagaat atgcagcgcc agagccgagg agaacccccg ctccctgagg 900
 aggacctgtc caaac tcttc aaaccaccac agccgcctgc caggatggac tcgctgtca 960
 ttgcaggcca gataa acact tactgccaga acatcaagga gttcactgcc caaacttag 1020
 gcaagctctt catggrccag gctcttcaag aatacaacaa ctaagaaaag gaa gtttcca 1080
 gaaaagaagt taaca tgaac tcttgaagtc acaccagggc aactcttgga agaataatat 1140
 ttgcatattg aaaag cacag aggatttctt tagtgtcatt gccgattttg gctataacag 1200
 tgtcttttcta gccataataa aataaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
 aaaaaaaaaa aaaaa aaaaa 1280

<210> 130
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003756

<400> 130
 tgagccaaga tatagttaaa tacaacacat acatgaggaa tactagtaaa caacagcagc 60

<210> 131
 <211> 839
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003832

<400> 131
 aagccacagg ctccc tgggt ggcgtcagct aaagtggctg ttgggtgtcc gcaggcttct 60
 gcctggccgc cgccgcctat aagctaccag gaggagcttt acgacttccc gtcctgcggg 120
 aagtggcggg cacga tcgca aggtagcgca gaagcttctc aatggccagc gccagctgca 180
 gcccggcgcg cgcactcgcc tcacctgagc ctgggaggaa aattcttcca aggatgatct 240
 cccactcaga gctga ggaag cttttctact cagcagatgc tgtgtgtttt gatgttgaca 300
 gcacggctcat cagtgaagaa ggaatcggat gctttcattg gatttggagg aaatgtgatc 360
 aggcaacaag tcaaggataa cgccaaatgg tatatcactg attttgtaga gctgctggga 420

```

gaaccggaag aaataacatcc attgtcatcac agctccaaac aacttcagat gaattttttac 480
aagttacaca gattgatact gtttgcttac aattgcctat tacaacttgc tataaaaagt 540
tggtacagat gatctgcact gtcaagtaaa ctacagttag gaatcctcaa agattgggttt 600
gtttgttttt aa.ctgtagtt ccagttattat atgatcacta tcgattttcct gga.gagtttt 660
gtaatctgaa tt.ctttatgt atattcctag ctatattttca tacaaagtgt ttt.aagagtg 720
gagagtcaat taaacacctt tactcttagg aatatagatt cggcagcctt cagtgaatat 780
tggttttttt cc.cttttgga tgtcaataaa agttttatcca tgtgtcagaa aaaaaaaaaa 839

```

<210> 132

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 132

```

gaagaaggaa tcggatgctt tcattggatt tggaggaaat gtgatcaggc aac.aagtcaa 60

```

<210> 133

<211> 3128

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 133

```

gcttcgcccc gtggcgcggt ttgaaatttt gcggggctca acggctcgcg gagcggctac 60
gcggagtgac atcgccggtg tttgcgggtg gttgttgctc tcggggccgt gtggagtagg 120
tctggacctg gactcacggc tgccttgagc gtccgccatg aggagaagtg aggtgctggc 180
ggaggagtcc atagtatgtc tgcagaaagc cctaaatcac cttcgggaaa tatgggagct 240
aattgggatt ccagaggacc agcgggtaca aagaactgag gtggtaaaga agc.atatcaa 300
ggaactcctg gat.tatgatga ttgctgaaga ggaaagcctg aaggaaagac tca.tcaaaag 360
catatccgtc tgtcagaaag agctgaacac tctgtgcagc gagttacatg ttgagccatt 420
tcaggaagaa ggagagacga ccatcttgca actagaaaaa gatttgcgca ccc.aagtgga 480
attgatgcga aaacagaaaa aggagagaaa acaggaactg aagctacttc aagagcaaga 540
tcaagaactg tgcgaaattc tttgtatgcc ccactatgat attgacagtg cct.cagtgcc 600
cagcttagaa gagctgaacc agttcaggca acatgtgaca actttgaggg aaa.caaaggc 660
ttctaggcgt gagggagtgtg tcagtataaa gagacagatc atactgtgta tggaaagaatt 720
agaccacacc ccagacacaa gctttgaaag agatgtgggtg tgtgaagacg aagatgcctt 780
ttgtttgtct ttggagaata ttgcaacact acaaaagtgt ctacggcagc tggaaatgca 840
gaaatcacaa aatgaagcag tgtgtgaggg gctgcgtact caaatccgag agc.tctggga 900
caggttgcaa atacctgaag aagaaagaga agctgtggcc accattatgt ctgggtcaaaa 960
ggccaaggtc cggaaagcgc tgcaattaga agtggatcgg ttggaagaac tga.aaatgca 1020
aaacatgaag aaagtgattg aggcaattcg agtggagctg gttcagtact gggaccagtg 1080
cttttatagc caggagcaga gacaagcttt tgcccctttc tgtgctgagg act.acacaga 1140
aagtctgtc ca.gctccacg atgctgagat tgtgcggtta aaaaactact atgaagttca 1200
caaggaactc tt.tgaagggtg tccagaagtg ggaagaaacc tggaggcctt tct.tagagtt 1260
tgagagaaaa gcttcagatc caaatcgatt taciaaccga ggaggaaatc ttt.taaaaga 1320
agaaaaacaa cgaagccaagc tccagaaaat gctgcccag ctggaagaag agt.tgaaggc 1380
acgaattgaa ttgtgggaac aggaacattc aaaggcattt atggtgaatg ggcagaaatt 1440
catggagtat gtggcagaac aatgggagat gcacgattg gagaaagaga gag.ccaagca 1500
ggaaagacaa ctgaagaaca aaaaacagac agagacagag atgctgtatg gcagcgctcc 1560
tcgaacacct agcaagcggc gaggactggc tcccaatata ccgggcaaag cacgtaagct 1620
gaacactacc accatgtcca atgctacggc caatagtagc attcggccta tct.ttgagg 1680
gacagtctac ca.ctcccccg tgtctcgact tcctccttct ggcagcaagc cagt.cgctgc 1740
ttccacctgt tcagggaaga aaacaccccc tactggcagg catggagcca aca.aggagaa 1800
cctggagctc aa.cggcagca tcctgagtgg tgggtaccct ggctcggccc ccc.tccagcg 1860
caacttcagc at.taattctg ttgccagcac ctattctgag tttgcgaagg atc.cgtccct 1920

```

ctctgacagt	tcca	ctggtt	ggcttcagcg	agaactttca	aaggcttcca	aa	tctgatgc	1980
tactttctgga	atcct	caatt	caaccaacat	ccagtcctga	gaagccctga	tc	agtcaccc	2040
agctgtggct	tcct	gtgcct	agactggacc	taattatatg	ggggtgactt	ta	gttttttct	2100
tcagcttagg	cgtg	cttgaa	accttggcca	ggttccatga	ccatgggcct	aa	cttaaaga	2160

tgtgaatgag	tgtta	cagtt	gaaagcccat	cataggttta	gtggtcctag	ga	gacttgg	2220
tttgacttat	atacat	gaaa	agtttatggc	aagaagtgc	aatttttagca	ta	tggggcct	2280
gactttctcta	ccacata	aatt	ctacttgctg	aagcatgatc	aaagcttggt	tt	atttcacc	2340
actgtaggaa	aatga	attgac	tatgcccatc	cctgggggta	attttggcat	gt	atacctgt	2400
aactagtaat	taacat	ctttt	tttgttttagg	catgttcaat	taatgctgta	gc	tatcatag	2460
ctttgctctt	acct	gaagcc	ttgtcccccac	cacacaggac	agccttcctc	ct	gaagagaa	2520
tgtctttgtg	tgtc	gaagt	tgagatggcc	tgccctactg	ccaaagaggt	ga	caggaagg	2580
ctgggagcag	cttt	gttaaa	ttgtgttcag	ttctgttaca	cagtgcattg	cc	ctttgttg	2640
ggggtatgca	tgtat	gaaca	cacatgcttg	tccgaacgct	ttctcggcgt	tt	gtcccttg	2700
gctctcatct	cccc	attcc	tgtgcctact	ttgcctgagt	tcttctaccc	cc	gcagttgc	2760
cagccacatt	gggag	tctgt	ttgttccaat	gggttgagct	gtctttgtog	tg	gagatctg	2820
gaactttgca	catgt	caacta	ctggggaggt	gttcctgctc	tagcttccac	ga	tgaggcgc	2880
cctctttacc	tatc	ctctca	atcactactc	ttcttgaagc	actattatth	at	tcttcgc	2940
tgtctgcctg	cagcag	tact	actgtcaaca	tagtgtaaat	ggttctcaaa	ag	cttaccag	3000
tgtggacttg	gtgtt	agcca	cgctgtttac	tcatacagta	cgtgtcctgt	tt	ttaaaata	3060
tacaattatt	ctta	aaaaata	aattaaaatc	tgtatactta	catttcaaaa	ag	aaaaaaaa	3120
aaaaaaaa		3128						

<210> 134
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003981

<400> 134	
tgcagcagta	ctactgtcaa catagtgtaa atggttctca aaagcttacc agtgtggact 60

<210> 135
 <211> 1816
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004029

<400> 135					
ggcaccaggg	gtccggcctg cgccttcccg ccaggcctgg aactgggttc aa	cacctgtg	60		
acttcatgtg	tgcgcgcggg ccacacctgc agtcacacct gtagccccc	ct	gccaagag 120		
atccataccg	aggcagcgtc ggtggctaca agccctcagt ccacacctgt	gg	acacctgt 180		
gacacctggc	cacacgacct gtggccggcg cctggcgtct	gctgcgacag	ga	gcccttac 240	
ctccccgtgt	ataaacacct accgccacct aactgcccct	gcagaaggag	ca	atggcctt 300	
ggctcctgag	agggcagccc cagcgtgct gttcggagag	tggctccttg	ga	gagatcag 360	
cagcggctgc	tatgagggggc tgcagtggct	ggacgaggcc	cgcacctgtt	tc	cgctgcc 420
ctggaagcac	ttcgcgcgca aggacctgag	cgaggccgac	gcgcgcctct	tc	aaggcctg 480
ggctgtggcc	cgcggcaggt ggccgcctag	cagcagggga	ggtggcccgc	cc	cccagggc 540
tgagactgcg	gagcgcggcg gctggaaaac	caacttccgc	tgcgcactgc	gc	agcacgcg 600
tcgcttctgt	atgctgcggg ataactcggg	ggacccggcc	gaccgcaca	ag	gtgtacgc 660
gctcagccgg	gagctgtgct ggcgagaagg	cccaggcacg	gaccagactg	ag	gcagaggc 720
ccccgcagct	gtcccaccac cacagggtgg	gccccagggg	ccattcttgg	ca	cacacaca 780
tgtgtgactc	caagccccag gccccctccc	tgtcccagct	ggtgacaagg	gg	gacctcct 840
gctccaggca	gtgcaacaga gctgcctggc	agaccatctg	ctgacagcgt	ca	tggggggc 900
agatccagtc	ccaaccaagg ctcttgaga	gggacaagaa	gggcttcccc	tg	actggggc 960
ctgtgctgga	ggcgaggccg cggccccaga	gtccccgcac	caggcagagc	cg	tacctgtc 1020
accctcccca	agcgcctgca ccgcgggtga	agagcccagc	ccagggggcg	tg	gacgtgac 1080

catcatgtac	aagggccgca	cgggtgctgca	gaaggtggtg	ggacacccga	gc tgcacgtt	1140
cctatacggc	ccccagacc	cagctgtccg	ggccacagac	ccccagcagg	tagcattccc	1200
cagccctgcc	gagctcccgg	accagaagca	gctgcgctac	acggagggaac	tgcctgctggca	1260
cgtggccctt	gggttgccac	tggagcttcg	ggggccacag	ctgtggggccc	ggcgcatggg	1320
caagtgcgaag	gtgtaactggg	aggtggggcgg	acccccaggc	tccgccagcc	cttccacccc	1380
agcctgcctg	ctgcctcgga	actgtgacac	ccccatcttc	gacttcagag	tccttcttcca	1440
agagctgggtg	gaattccggg	cacggcagcg	cgtgggtcc	ccacgctata	ccatctacct	1500
gggcttcggg	caggacctgt	cagctgggag	gcccaggag	aagagcctgg	tcctggtgaa	1560
gctggaaccc	tggctgtgcc	gagtgacact	agagggcacg	cagcgtgagg	gtgtgtcttc	1620
cctggatagc	agcagcctca	gcctctgcct	gtccagcgcc	aacagcctct	atgacgacat	1680
cgagtgtctc	cttatggagc	tggagcagcc	cgcctagaac	ccagtctaata	gagaactcca	1740
gaaagctgga	gcagccacc	tagagctggc	cgcggccgcc	cagtctaata	aaaagaactc	1800
cagaacaaaa	aaaaaa	1816				

<210> 136

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004029

<400> 136

agcagccac ctagagctgg ccgcgccgc ccagtctaata aaaaagaact ccagaacaaa 60

<210> 137

<211> 2121

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004203

<400> 137

tggaatTTTT	ggcgcgagca	gctccgcgcg	cgttcacggg	ccgttcccc	tcacggggagt	60
cctccgccc	ggcgtccgga	acagtcgacg	gcagactccg	gcccgtgag	ccacccgagg	120
gggtcccggtg	cctccgcgga	cccggaaatct	gggcccctgc	ggacccgcgc	cccgcccgagt	180
cgccccagg	cttcccaaca	cccacggagt	gaagtcagcc	gcggccctgc	ctggggaggaa	240
cttaccgtct	accgggaaag	gtggccagca	gatgtgtcgg	gcctggtgag	aggggtgaggc	300
gagacggccc	gatcgccag	ggccccggaa	gctgcggagg	tcacccccgc	ctgggccttag	360
ctcagggaca	ccctggattc	acgtgggagc	ccctgctcct	gcctcccccg	tcccaccact	420
gaggctgttg	ggccaggcca	gtcatgctag	aacggcctcc	tgcactggcc	atgcccctgc	480
ccacggagg	cacccgcca	cctctgagtg	gcacccccat	ccagtccca	gcctacttcc	540
gccacgcaga	acctggattc	tccctcaaga	ggccccagg	gctcagccgg	agcctccac	600
ctccgcccc	tgccaaaggc	agcattccca	tcagccgcct	cttccctcct	cggaccccag	660
gctggcacca	gctgcagccc	cggcgggtgt	cattccgggg	cgaggcctca	gagactctgc	720
agagccctgg	gtatgaccca	agccggccag	agtccttctt	ccagcagagc	ttccagaggc	780
tcagccgcct	gggcctatgg	tcctacggag	aggtcttcaa	ggtgcgctcc	aaggaggacg	840
gccggctcta	tgcggtaaag	cgttccatgt	caccattccg	gggccccaa	gaccggggccc	900
gcaagtggg	cgaggtgggc	agccacgaga	aggtggggca	gcacccatgc	tgcgtgcggc	960
tggagcaggc	ctgggaggag	ggcggcatcc	tgtacctgca	gacggagctg	tgcggggcca	1020

gcctgcagca	acactgtgag	gcctgggggtg	ccagcctgcc	tgaggcccag	gtctgggggt	1080
acctgcggga	cacgctgctt	gccctggccc	atctgcacag	ccagggcctg	gtgcaccttg	1140
atgtcaagcc	tgccaacatc	ttcctggggc	cccggggccg	ctgcaagctg	ggtgacttcg	1200
gactgctggt	ggagctgggt	acagcaggag	ctggtgaggt	ccaggaggga	gacccccgct	1260
acatggcccc	cgagctgctg	cagggctcct	atgggacagc	agcggatgtg	ttcagctctg	1320
gcctcaccat	cctggaagtg	gcattgcaaca	tggagctgcc	ccacggtggg	gagggctggc	1380
agcagctgcg	ccagggtctac	ctgccccctg	agttcactgc	cggctctgtct	tcagagctgc	1440
gttctgtcct	tgtcatgatg	ctggagccag	acccaagct	gcgggcccag	gccgaggccc	1500
tgtctggcact	gcctgtgttg	aggcagccgc	gggcctgggg	tgtgctgtgg	tgcattggcag	1560

cgaggagccct	gagccgaggg	tgggcccctgt	ggcaggccct	gcttgcccctg	ctctgctggc	1620
tctggcatgg	gctggctcac	cctgccagct	ggctacagcc	cctggggcccg	ccagccaccc	1680
cgcttggtc	accaccctgc	agtttgctcc	tggacagcag	cctctccagc	aactgggatg	1740
acgacagcct	agggccttca	ctctcccctg	aggctgtcct	ggcccggact	gtgggggagca	1800
cctccacccc	ccggagcagg	tgcacaccca	gggatgccct	ggacctaaagt	gacatcaact	1860
cagagcctcc	tcgggggctcc	ttcccctcct	ttgagcctcg	gaacctcctc	agcctgtttg	1920
aggacaccct	agaccaaac	tgagccccag	actctgcctc	tgcactttta	accttttatc	1980
ctgtgtctct	cccgtcgccc	ttgaaagctg	gggcccctcg	ggaactocca	tggctcttctc	2040
tgcctggccg	tgtctaataa	aaagtatttg	aaccttggga	gcacccaagc	ttgctcatgt	2100
ggcaaaaaaa	aaaaaaaaaa	a	2121			

<210> 138

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_O04203

<400> 138

ctggccgtgt ctaataaaaa gtatttgaac cttgggagca cccaagcttg ctcatgtggc 60

<210> 139

<211> 1982

<212> DNA

<213> Homo sapiens

<300>

<308> NM_O04207

<400> 139

ggcgagaggc	gggctgaggg	ggcccagcgg	cggcaggtga	ggcggaacca	accctcctgg	60
ccatgggagg	ggcgtgggtg	gacgagggcc	ccacaggcgt	caaggcccct	gacggcggtc	120
ggggctgggc	cgtgctcttc	ggctgtttcg	tcactactgg	cttctcctac	gccttcccca	180
aggcgtcag	tgtcttcttc	aaggagctca	tacaggagtt	tgggatcggc	tacagcgaca	240
cagcctggat	ctcctccatc	ctgctggcca	tgtctacgg	gacagggtccg	ctctgcagt	300
tgtgcgtgaa	ccgctttggc	tgccggcccc	tcattgcttg	gggggggtctc	tttgctctgc	360
tgggcatggt	ggctgcgtcc	ttttgccgga	gcactcatca	ggtctacctc	accactgggg	420
tcactacggg	gttgggtttg	gcactcaact	tccagccctc	gctcatcatg	ctgaaccgct	480
acttcagcaa	gcggcgcccc	atggccaacg	ggctggcgcc	agcaggtagc	cctgtcttcc	540
tgtgtgccct	gagcccgcgtg	gggcagctgc	tgcaggaccg	ctacggctgg	cggggcggtc	600
tcttcatect	ggggcgccctg	ctgctcaact	gctgcgtgtg	tgccgcactc	atgaggcccc	660
tgggtgtcac	ggcccagccg	ggctcggggc	cgccgcgacc	ctcccggcgc	ctgctagacc	720
tgagcgtctt	ccgggacccg	ggctttgtgc	tttacgcctg	ggccgcctcg	gtcatgggtc	780
tggggctctt	cgtcccgcgc	gtgttcgtgg	tgagctacgc	caaggacctg	ggcgtgccc	840
acaccaaggc	cgccttctctg	ctcaccatcc	tgggttcat	tgacatcttc	gcgcggccgg	900
ccgcgggctt	cgtggcgggg	cttgggaagg	tgcggcccta	ctccgtctac	ctcttcagct	960
tctccatgtt	cttcaacggc	ctcgcgacc	tggcgggctc	tacggcgggc	gactacggcg	1020
gcctcgtggt	cttctgcata	ttctttggca	tctcctacgg	catggtgggg	gccctgcagt	1080
tcgaggtgct	catggccatc	gtgggcaccc	acaagttctc	cagtgccatt	ggcctgggtc	1140
tgtgatgga	ggcgggtggc	gtgctcgtcg	ggcccccttc	gggaggcaaa	ctcctggatg	1200
cgaccacgct	ctacatgtac	gtgttcaccc	tggcgggggc	cgagggtgctc	acctctctcc	1260
tgattttgct	gctgggcaac	ttcttctgca	ttagggaagaa	gcccacagag	ccacagcctg	1320
agggtggcgg	cgcgaggagg	gagaagctcc	acaagcctcc	tgcagactcg	gggggtggact	1380
tgcgggagggt	ggagcatttc	ctgaaggctg	agcctgagaa	aaacggggag	gtgggttcaca	1440
ccccggaaa	aagtgtctga	gtggctgggc	ggggccggca	ggcacaggga	ggaggtacag	1500
aagccggcaa	cgcttgctat	ttattttaca	aactggactg	gctcaggcag	ggccacggct	1560
gggctccagc	tgccggccca	gcggatcgct	gcccgatcag	tgttttgagg	gggaagggtg	1620
cgggggtggga	accgtgtcat	tccagagtgg	atctgcgggtg	aagccaagcc	gcaagggttac	1680
aaggcatcct	caccaggggc	cccgcctgct	gctcccagggt	ggcctgcggg	cactgctatg	1740
ctcaaggac	tggaaaccca	tgcttcgaga	caacgtgact	ttaatggggag	ggtgggtggg	1800

```

ccgcagacag gctggcaggg caggtgctgc gtggggccct ctccagcccg tccctacctg 1860
ggctcacatg gggcctgtgc ccacccctct tgagtgtctt ggggacagct ctttccaccc 1920
ctggaagatg gaaataaacc tgcgtgtggg tggagtgttc tcgtgccgaa ttcaaaaagc 1980
tt 1982

```

<210> 140

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004207

<400> 140

```

cctcttgagt gtcttgggga cagctctttc cacccttgga agatggaaat aaacctgcgt 60

```

<210> 141

<211> 2054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004209

<400> 141

```

cgggaggcgg cagcggctgc agcgttggta gcatcagcat cagcatcagc ggcagcggca 60
gcggcctcgg gcggggccgg ccggacggac aggcggacag aaggcgccag gggcgcgcg 120
cccgcccggg ccggccatgg agggcgccct cttcggcgcg ggccgcgcag gggcgcgcc 180
ggaccccgctg agctttgcgc ggcgggccca gaccctgtct cgggtcgcgt cctgggtgtt 240
ctccatcgcc gtcttcgggc ccatcgctca cgagggtac gtgaacaccg acagcggccc 300
cgagctgcgc tgcgtgttca acgggaacgc gggcgccctgc cgcttcggcg tcgcgtggg 360
cctcggagcc ttcctcgcc tgcgcgcctt cctgctgtct gatgtgcgt tcagcaaat 420
cagcagcgtc cgcgaccgcc ggcgcgcggt gttgctggac ctgggcttct caggactctg 480
gtccttcctg tggttcgtgg gcttctgtct cctaccaat cagtggcagc gcacggcgcc 540
agggccggcc acgacgcagg cgggggacgc ggcgcgggcc gccatcgct tcagcttctt 600
ctccatcctc agctgggtgg cgctcacgt gaaggccctg cagcggttcc gctgggcac 660
cgacatgtca ctcttcgcca ccgaacagct gagcaccggg gcgagccagg cctaccccg 720
ctatccggtg ggcagcggcg tggagggcac cgagacctac cagagccgc ccttcaccga 780
gacctggac accagcccca aagggtacca ggtgcccgcc tactagcggc tggcaggcac 840
agaccagggc tccaaggcca cccaccaaac gcaggcccca gggctctcgg gacctccctt 900
gggtccttcc agctcagtg cgcggacaga gtaggtggcc gctttgcgc atccggggcc 960
aagagggggg ggaaccgcgt gtctgggctg cccctgccaa gttccccag tcctcagca 1020
cctggcccca ggaactgagg cctgagaagg ggaatagcact gccaggacg tgtgtcccta 1080
gctggaatg gactggcctg gggaaggctt tccctcttg ggccacacct gctcactctg 1140
gggttggggg tccagctgcc ctctacgata aggtgcaggg gctgccagg acaagcggg 1200
ggcaggggaa agacaccacc ctgcgcccaa gactggggat cctggccact gtcccatcc 1260
catgtccctg tgggtagtga ctgtctcgtt tctgtcatgg tgggtgcgtc cgtccggagc 1320
cactctccac tttctctcac aggtgtctag aacagcccag ccctgtcagt gtgtgatca 1380
tgggtccagt ttcgggtttc acctcctagt actccacaag ctgctcctct ctctgtggcc 1440
ccggccccctg ccaggtgtg ggtggttctg gccaggaagg cacaaggtag ctgtgggcca 1500
agacaccagc cctgtcctag cccttcagta agaccttgcc aggagaggag aaggatgct 1560
gggtgccagg caagacaagc cctcagcag gagagaggcc cagaggctcc agctggccac 1620
cgtgccccac aagatggccc ctgtgtgggt ccccttacct tggcttcctg gccagtcct 1680
tgcccttcca cctgcaccct gcttcctggc ccagtcccag gttggagtcc ctctgcatag 1740
ctgactactc atgcattgct caaagctggc ttttcacatt aagtcaacac caaacgtgg 1800
tgccacattt catcagacag acacctccct ctggagatgc agttgagtga caaccttg 1860
acattgtagc ctagaccaat tctgtgtgga tatttaagt aacatgttta caattttgt 1920
atatatcact ctctccctct cctgaaagac cagagattgt gtattttcag tgtcccatgt 1980
tcgactgca cctctctttac aataaagact gtaactgagc tgactgtgaa aaaaaaaaa 2040
aaaaaaaaa aaaa 2054

```

<210> 142
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004209

<400> 142
 gatgcagttg agtgacaacc ttgttacatt gtagcctaga ccaattctgt gtggatattt 60

<210> 143
 <211> 1224
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 143
 ggccgggaga gtagcagtgc cttggacccc agctctcctc cccctttctc tctaaggatg 60
 gccagaagg agaactccta cccctggccc tacggccgac agacggctcc atctggcctg 120
 agcaccctgc ccagcgagt cctccggaaa gagcctgtca ccccatctgc acttgtcctc 180
 atgagccgct ccaatgtcca gccacagct gccctggcc agaaggatg ggagaatagc 240
 agtgggacac ccgacatctt aacgcggcac ttcacaattg atgactttga gattgggcgt 300
 cctctgggca aaggcaagtt tggaaacgtg tacttggctc gggagaagaa aagccatttc 360
 atcgtggcgc tcaaggctct cttcaagtcc cagatagaga aggagggcgt ggagcatcag 420
 ctgcgagag agatcgaaat ccaggcccac ctgcaccatc ccaacatcct gcgtctctac 480
 aactattttt atgaccggag gaggatctac ttgattctag agtatgcccc ccgcggggag 540
 ctctacaagg agctgcagaa gagctgcaca ttgacgagc agcgaacagc cacgatcatg 600
 gaggagtgg cagatgctct aatgtactgc catgggaaga aggtgattca cagagacata 660
 aagccagaaa atctgtctctt agggctcaag ggagagctga agattgctga cttcggctgg 720
 tctgtgcatg cgcctctcct gaggaggaag acaatgtgtg gcaccctgga ctacctgcc 780
 ccagagatga ttgaggggag catgcacaat gagaagggtg atctgtggtg cattggagtg 840
 ctttgctatg agctgctggg ggggaaccca cctttgaga gtgcatcaca caacgagacc 900
 tatcgccgca tcgtcaagggt ggacctaaag ttccccgctt ctgtgcccac gggagcccag 960
 gacctcatct ccaaactgct caggcataac ccctcggaac ggctgccccct ggcccaggctc 1020
 tcagcccacc cttgggtccg ggccaactct cggagggtgc tgccctccctc tgcccttcaa 1080
 tctgtgcct gatggctcct gtcattcact cgggtgctg tgtttgtatg tctgtgtatg 1140
 tataggggaa agaagggatc cctaactgtt cccttatctg ttttctacct cctcctttgt 1200
 ttaataaagg ctgaagcttt ttgt 1224

<210> 144
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 144
 gtctgtgtat gtatagggga aagaaggat ccctaactgt tcccttatct gttttctacc 60

<210> 145
 <211> 983
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004335

```

<400> 145
gtggaattca tggcatctac ttcgtatgac tattgcagag tgcccatgga agacggggat 60
aagcgtctgta agcttctgct ggggatagga attctggtgc tcctgatcat cgtgattctg 120
ggggtgccct tgattatctt caccatcaag gccaacagcg aggcctgccg ggacggcctt 180
cgggcagtgga tggagtgctg caatgtcacc catctcctgc aacaagagct gaccgaggcc 240
cagaaggggt ttcaggatgt ggaggccag gccgccacct gcaaccacac tgtgatggcc 300
ctaattggctt ccttggatgc agagaaggcc caaggacaaa agaaagtgga ggagcttgag 360
ggagagatca ctacattaaa ccataagctt caggacgctg ctgcagaggt ggagcgactg 420
agaagagaaa accaggtctt aagcgtgaga atcgcggaaca agaagtacta cccagctcc 480
caggactcca gctccgctgc ggcccccag ctgctgattg tgctgctggg cctcagcgct 540
ctgctgcagt gagatcccag gaagctggca catcttggaa ggtccgtcct gctcggttt 600
tcgcttgaaac attcccttga tctcatcagt tctgagcggg tcatggggca acacggttag 660
cggggagagc acggggtagc cggagaaggg cctctggagc aggtctggag gggccatggg 720
gcagtcctgg gtgtggggac acagtcgggt tgaccaggg ctgtctccct ccagagcctc 780
cctccggaca atgagtcctt cctcttgtct cccaccctga gattgggcat ggggtgcggt 840
gtggggggca tgtgctgcct gttgttatgg gttttttttg cggggggggg tgcctttttt 900
tggggtcttt gagctccaaa aaataaacac ttcctttgag ggagagcaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaa 983

```

```

<210> 146
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004335

```

```

<400> 146
ggttgctttt ttctgggggtc tttgagctcc aaaaaataaa cacttccttt gagggagagc 60

```

```

<210> 147
<211> 3446
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004336

```

```

<400> 147
ttctagtttg cggttcaggt ttgccgctgc cggccagcgt cctctggcca tggacacccc 60
ggaaaatgtc cttcagatgc ttgaagccca catgcagagc tacaagggca atgaccctct 120
tggtgaatgg gaaagataca tacagtgggt agaagagaat ttctctgaga ataaagaata 180
cttgataact ttactagaac atttaatgaa ggaattttta gataagaaga aataccacaa 240
tgaccaaga ttcatcagtt attgtttaaa atttgctgag tacaacagtg acctccatca 300
attttttgag tttctgtaca accatgggat tggaaacctg tcatcccctc tgtacattgc 360
ctgggctggg catctggaag cccaaggaga gctgcagcat gccagtgtct tcttccagag 420
aggaattcaa aaccaggctg aaccagaga gttcctgcaa caacaataca ggttatttca 480
gacacgcctc actgaaaccc atttgccagc tcaagctaga acctcagaac ctctgcataa 540
tggtcaggtt ttaaatcaaa tgataacatc aaaatcaaat ccaggaaata acatggcctg 600
catttctaag aatcagggtt cagagctttc tggagtata tcttcagctt gtgataaaga 660
gtcaaatatg gaacgaagag tgatcacgat ttctaaatca gaataattct tgcactcatc 720
tttggtcatc aaagttgatg ttgagcaggt tggtatgtat tgcaaggaga agcttattcg 780
tggggaatca gaattttcct ttgaagaatt gagagcccag aaatacaatc aacggagaaa 840
gcatgagcaa tgggtaaatg aagacagaca ttatatgaaa aggaaagaag caaatgcttt 900
tgaagaacag ctattaaaac agaaaatgga tgaacttcat aagaagttgc atcaggtggg 960
ggagacatcc catgaggatc tgcccgcctt ccaggaaagg tccgaggtta atccagcacg 1020
tatggggcca agtgtaggct cccagcagga actgagagcg ccatgtcttc cagtaacctc 1080
tcagcagaca ccagtgaaca tggaaaagaa cccaagagag gcacctcctg ttgttcctcc 1140
tttggcaaatt gctatttctg cagctttggt gtcccagcc accagccaga gcatttgctc 1200
tcctgttcct ttgaaagccc agacagtaac agactccatg tttgcagtgg ccagcaaaga 1260
tgctggatgt gtgaataaga gtactcatga attcaagcca cagagtggag cagagatcaa 1320

```

```

agaaggggtgt gaaacacata aggttgccaa cacaagttct tttca.cacaa ctccaaacac 1380
atcactggga atggttcagg caacgccatc caaagtgcag ccatc.accca ccgtgcacac 1440
aaaagaagca ttaggtttca tcatgaatat gtttcagggt cctac.acttc ctgatatttc 1500
tgatgacaaa gatgaatggc aatctctaga tcaaaatgaa gatgc.atttg aagcccagtt 1560
tcaaaaaaat gtaagggtcat ctggggcttg gggagtcagt aagat.catct cttctttgtc 1620
atctgctttt catgtgtttg aagatggaaa caaagaaaat tatgg.ATTAC cacagcctaa 1680
aaataaaacc acaggagcca ggacctttgg agaacgctct gtcag.cagac ttccttcaaa 1740
accaaaggag gaagtgcctc atgctgaaga gtttttggat gactc.aactg tatgggggat 1800
tcgctgcaac aaaaccctgg caccagtc taagagccca ggaga.cttca catctgctgc 1860
acaacttgcg tctacaccat tccacaagct tccagtggag tcagt.gcaca ttttagaaga 1920
taaagaaaat gtggttagcaa aacagtgtac ccaggcgact ttgga.tttct gtgaggaaaa 1980
catggtgggt ccttcaaggg atggaaaatt cagtccaatt caaga.gaaaa gcccaaaaca 2040
ggccttgctg tctcacatgt attcagcatc cttacttcgt ctgag.ccagc ctgctgcagg 2100
tggggtactt acctgtgagg cagagttggg cgttgagggt tgcag.actca cagacactga 2160
cgctgccatt gcagaagatc caccagatgc tattgtctgg ctcca.agcag aatggatgca 2220
gatgagttca cttgggactg ttgatgctcc aaacttcatt gttgg.gaacc catgggatga 2280
taagctgatt ttcaaacttt tatctgggct ttctaaacca gtgag.ttcct atccaaatac 2340
ttttgaatgg caatgtaaac ttccagccat caagcccaag actga.atttc aattgggttc 2400
taagctgggt tatgtccatc accttcttgg agaaggagcc tttgc.ccagg tgtacgaagc 2460
taccaggga gatctgaatg atgctaaaaa taaacagaaa tttgt.tttaa aggtccaaaa 2520
gcctgccaac ccctgggaat tctacattgg gaccagttg atgga.aagac taaagccatc 2580
tatgcagcac atgtttatga agttctattc tgcccactta ttcca.gaatg gcagtgtatt 2640
agtaggagag ctctacagct atggaacatt attaaatgcc attaa.cctct ataaaaatac 2700
ccctgaaaaa gtgatgcctc aaggtcttgt catctctttt gctat.gagaa tgctttacat 2760
gattgagcaa gtgcatgact gtgaaatcat tcatggagac attaa.accag acaatttcat 2820
acttggaac ggatttttgg aacaggatga tgaagatgat ttatc.tgctg gcttggcact 2880
gattgacctg ggtcagagta tagatatgaa actttttcca aaagg.aaacta tattcacagc 2940
aaagtgtgaa acatctgggt ttcagtgtgt tgagatgctc agcaa.caaac catggaacta 3000
ccagatcgat tactttgggg ttgctgcaac agtatattgc atgct.ctttg gcacttacat 3060
gaaagtgaaa aatgaaggag gagagtgtaa gcctgaagg cttaa.tagaa ggcttcctca 3120
tttgatgatg tggaatgaat tttttcatgt tatgttgaat attcc.agatt gtcatcatct 3180
tccatctttg gatttggtta ggcaaaagct gaagaaagta tttca.acaac actatactaa 3240
caagattagg gccctacgta ataggcta atgta.ctgctc ttaga.atgta agcggttcacg 3300
aaaataaaat ttggatatag acagtcctta aaaatcacac tgtaa.atatg aatctgctca 3360
ctttaaacct gttttttttt catttattgt ttatgtaaat gtttgr.ttaa aataaatccc 3420
atggaatatt tccatgtaaa aaaaaa 3446

```

<210> 148
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004336

<400> 148
 ttagggccct acgtaatagg ctaattgtac tgctcttaga atgta.agcgt tcacgaaaat 60

<210> 149
 <211> 739
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004345

<400> 149
 taaagcaaac cccagcccac accctggcag gcagccaggg atggg.tggat caggaaggct 60
 cctgggtggg ctttttgcac aggtc.caggc tgggcataaa ggagg.ctcct gtgggctaga 120
 gggaggcaga catggggacc atgaagacc aaagggatgg ccact.ccctg gggcggtgg 180
 cactgggtgt cctgctgctg ggcctgggtga tgccctctggc catca.ttgcc caggtcctca 240

```

gctacaagga agctgtgctt cgtgctatag atggcatcaa ccagcgggtcc tcggatgcta 300
acctctaccg cctcctggac ctggacccca ggccacgat ggatggggac ccagacacgc 360
caaagcctgt gagcttcaca gtgaaggaga cagtgtgccc caggacgaca cagcagtcac 420
cagaggattg tgacttcaag aaggacgggc tgggtgaagcg gtgtatgggg acagtgaccc 480
tcaaccaggc caggggctcc tttgacatca gttgtgataa ggataacaag agatttgccc 540
tgctgggtga tttcttccgg aaatctaaag agaagattgg caaagagttt aaaagaattg 600
tccagagaat caaggatttt ttgcggaatc ttgtaccag gacagagtc tagtgtgtgc 660
cctaccctgg ctcaggcttc tgggctctga gaaataaact atgagagcaa tttcaaaaaa 720
aaaaaaaaa aaaaaaaaaa 739

```

```

<210> 150
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004345

```

```

<400> 150
gcaaagagtt taaaagaatt gtccagagaa tcaaggattt tttgcggaat cttgtaccca 60

```

```

<210> 151
<211> 1432
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004577

```

```

<400> 151
gaggaaaatt cttccagcga tggctctccca ctcagagctg aggaagcttt tctactcagc 60
agatgctgtg tgttttgatg ttgacagcac ggtcatcaga gaagaaggaa tcgatgagct 120
agccaaaatc tgtggcggtg aggacgcggt gtcagaaaatg acacggcgag ccatgggagg 180
ggcagtgctt tcaaaagctg ctctcacaga gcgcttagcc ctcattccagc cctccaggga 240
gcaggtgcag agactcatag cagagcaacc cccacacctg acccccgga taaggagct 300
ggtaagtgcg ctacaggagc gaaatgttca ggttttccca atatctggtg gctttaggag 360
tattgtagag catgttgctt caaagctcaa tatcccagca accaatgtat ttgccaatag 420
gctgaaattc tactttaacg gtgaatatgc aggttttgat gagacgcagc caacagctga 480
atctggtgga aaaggaaaag tgattaaact tttaaaggaa aaatttcatt ttaagaaaat 540
aatcatgatt ggagatgggtg ccacagatat ggaagcctgt cctcctgctg atgctttcat 600
tggatttgga ggaaatgtga tcaggcaaca agtcaaggat aacgccaaat ggtatatcac 660
tgattttgta gagctgctgg gagaactgga agaataacat ccattgtcgt acagctccaa 720
acaacttcag atgaattttt acaagttata cagattgata ctgtttgctt acagttgct 780
attacaactt gctatagaaa gttggtacaa atgatctgta ctttaaaacta cagttaggaa 840
tcctagaaga ttgctttttt ttttttttta actgtagtgc cagtattata tgatgactat 900
tgatttctctg gagagggtttt tttttttttt gagacagaat cttgctctgt tggccaggct 960
ggagtgcagt ggcgcggtct cggctcactg caagctctgc ctcccagggt cagccattc 1020
tcctgctca gcctcccag tagctgggac tacaggcacc cgccaccaca tccggcta 1080
tttttgtatt ttagtagag acgggggttg accgtgttag ccaggatggc cttgatctcc 1140
tgaccttggtg atccgcctgc ctcagcctcc caaagtgtg ggattacagg cttgggccac 1200
cgccccagc caatgtccta gagagttttg tgatctgaat tctttatgta tttttagc 1260
tatattcat acaaagtgct ttaagtgtgg agagtcaatt aaacacctt actcttagaa 1320
atacggatc ggcagccttc agtgaatatt ggtttctctt tggatatgtc aaaaaagttt 1380
atccgtatgt cagaacggat ttgtggaaaa aaaaaaaaaa aaaaaaaaaa aa 1432

```

```

<210> 152
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>

```

<308> NM_004577

<400> 152

tagaaatacg gattcggcag ccttcagtga atattgggtt ctctttggta tgtcaataaa 60

<210> 153

<211> 1530

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 153

```

aatcctggaa caaggctaca gcgtcgaaga tccccagcgc tgcgggctcg gagagcagtc 60
ctaacggcgc ctcgtaagct agtgtcctcc cttttcagtc cgcgtccctc cctggggcgg 120
gctggcactc ttgccttccc cgtccctcat ggcgtgtctc cgacgcccg aaggtgtccag 180
tgatttggag aatattgaca caggagttaa ttctaaagtt aagagtcattg tgactattag 240
gcgaactgtt ttagaagaaa ttggaaatag agttacaacc agagcagcac aagtagctaa 300
gaaagctcag aacaccaaag ttccagttca acccaccaaa acaacaaatg tcaacaaaca 360
actgaaacct actgcttctg tcaaaccagt acagatggaa aagttggctc caaaggggtcc 420
ttctcccaca cctgaggatg tctccatgaa ggaagagaat ctctgccaa gcttttctga 480
tgccttgctc tgcaaaatcg aggacattga taacgaagat tgggagaaac ctcagctctg 540
cagtgaactac gtttaaggata tctatcagta tctcaggcag ctggagggtt tgcagtcctat 600
aaaccacat ttcttagatg gaagagatat aaatggacgc atgctgtcca tcctagtggg 660
ttggctggta caagtccact ccaagtttag gcttctgcag gagactctgt acatgtgcgt 720
tggcattatg gatcgatttt tacaggttca gccagtttcc cggaagaagc ttcaattagt 780
tgggattact gctctgctct tggcttccaa gtatgaggag atgttttctc caaatattga 840
agactttgtt tacatcacag acaatgctta taccagttcc caaatccgag aaatggaaac 900
tctaattttg aaagaattga aatttgagtt gggctgaccc ttgccactac acttcttaag 960
gcgagcatca aaagccgggg aggttgatgt tgaacagcac actttagcca agtatttgat 1020
ggagctgact ctcatcgact atgatatggt gcattatcat ccttctaagg tagcagcagc 1080
tgcttctctg ttgtctcaga aggttctagg acaaggaaaa tggaaactta agcagcagta 1140
ttacacagga tacacagaga atgaagtatt ggaagtcatt cagcacatgg ccaagaatgt 1200
ggtgaaagta aatgaaaact taactaaatt catcgccatc aagaataagt atgcaagcag 1260
caaactcctg aagatcagca tgatccctca gctgaactca aaagccgtca aagaccttgc 1320
ctccccactg ataggaaggt cctaggtctg cgtggggccct ggggatgtgt gcttcattgt 1380
gccctttttc ttattgggtt agaactcttg attttgtaca tagtctcttg gtctatctca 1440
tgaaacctct tctcagacca gttttctaaa catatattga ggaaaaataa agcgattggg 1500
ttttcttaag gtaaaaaaaaa aaaaaaaaaa 1530

```

<210> 154

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 154

agaactcttg attttgtaca tagtctcttg gtctatctca tgaaacctct tctcagacca 60

<210> 155

<211> 2536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 155

agcgggtg	cg	60
gcccagcgt	ggtc	120
aagtagccgt	ttacaagcta	180
agcccagata	atccaggcca	240
taaacatggt	aaaaaaggag	300
ctgacttgga	accacagatg	360
tatacacact	tcatagggaa	420
tgacacaaaa	ggatataaat	480
ttgcttccaa	acttgaggaa	540
atggtgcttg	cagtgaagag	600
aatgggaact	ttgtcctgta	660
ctcttaaaga	tgctcctaaa	720
tagctcagct	tttagatctg	780
tactgactgc	tgctgccttg	840
gtttggagtg	ggacagtatt	900
taaaaagtac	tagtccagtg	960
ataatatcca	gacacataca	1020
ccttcagaaa	aggggggacag	1080
agagcactga	aaaaccacca	1140
ttcaccaaga	ttgggtagaa	1200
ctgtgattga	ttgccctagc	1260
caattggcac	taaagaatac	1320
gacttggtta	caaagatgtc	1380
ataccatagc	aatgtttttc	1440
aattttgtca	ctggagtttt	1500
acagtacttt	gaatctagtt	1560
ttggttcttg	aaaattaaat	1620
ctaattttatc	tatagctgct	1680
aaattttaatg	tttttttaaac	1740
aggcttagat	gaacatggtg	1800
ataaacaaga	taaatgctgt	1860
ttggtaataa	ccttttttga	1920
tatttatatg	aggaagtcac	1980
athtagtggtg	actaaggctt	2040
ctccattgtg	agataaggac	2100
tattttctta	aggaagaaga	2160
aaaaagacat	gaaaacttgc	2220
tttctaaaaa	aaaaatccag	2280
ttgttatcag	atatgtgaat	2340
tatgaagtct	tctgtatagt	2400
aagaaaacac	cactaatgtg	2460
tatttttagc	ccattgttta	2520
aaaaaaaaaa	aaaaaa	2536

<210> 156
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004702

<400> 156
 gtttgtgaaa ctgttaagggt cttttctaaa ttctccatt gtgagataag gacagtgtca 60

<210> 157
 <211> 1491
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_004710

<400> 157

```

gcggcgccg cagcgccggc gacggcgaca tggagagcgg ggcctacggc ggcggccaagg 60
cgggcggctc cttecgacct cggcgcttcc tgacgcagcc gcaggtgggt ggcgcgcgcc 120
tgtgcttggc cttegccttg atcgtgttct cctgcatacta tggtagagggc tacagcaatg 180
cccacgagtc taagcagatg tactgcgtgt tcaaccgcaa cgaggatgcc tgccgctatg 240
gcagtggcat cggggtgctg gccttcctgg cctcggcctt cttcttgggt gtcgacgcgt 300
atttcccca gatcagcaac gccactgacc gcaagtacct ggtcattggt gacctgctct 360
tctcagctct ctggaccttc ctgtggtttg ttggtttctg cttcctcacc aaccagtggg 420
cagtcacca cccgaaggac gtgctgggtg gggccgactc tgtgagggca gccatcacct 480
tcagcttctt ttccatcttc tcttgggtg tgctggcctc cctggcctac cagcgctaca 540
aggctggcgt ggacgacttc atccagaatt acgttgacct cactccggac cccaacactg 600
cctacgcctc ctaccagggt gcattctgtg acaactacca acagccacct ttcaccaga 660
acgcggaga caccgagggc taccagccgc cccctgtgta ctgagcggcg gttagcgtgg 720
gaaggggga agagagggcc ctcccctctg ccttggaact tcccatgagc ctctggaac 780
tgccagccc tctctttcac ctgttccatc ctgtgcagct gacacacagc taaggagcct 840
catagcctgg cgggggctgg cagagccaca cccaagtgc ctgtgccag agggcttcag 900
tcagccgctc actcctccag ggcattttta ggaaagggtt ttcagctagt gtttttctc 960
gcttttaata acctcagccc cgcctgcagt ggctagaagc cagcaggtgc ccatgtgcta 1020
ctgacaagt cctcagcttc ccccggccc gggctcaggc gtgggagccg ctattatctg 1080
cgttctctg caaagactcg tgggggcat cacacctgcc ctgtgcagcg gagccggacc 1140
aggctcttgt gtccctactc aggtttgctt cccctgtgcc cactgctgta tgatctgggg 1200
gccaccacc tgtgcgggtg gcctctgggc tgccctccgt ggtgtgaggg cggggctggt 1260
gctcatggca ttctctcctt gctcccacc ctggcagcag ggaagggtt tgcctgacaa 1320
caccagctt ctgttaata ttctgcagtt ttacttagg aagcctgggg agggcagggg 1380
tgcccatgg ctcccagact ctgtctgtgc cgagtgtatt ataaaatcgt gggggagatg 1440
ccgggcctgg gatgctgttt ggagacggaa taaatgtttt ctcatcagt a 1491

```

<210> 158

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 158

```

ttgcctgaca acaccagct ttatgtaaat attctgcagt tgttacttag gaagcctggg 60

```

<210> 159

<211> 332 4

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 159

```

gcagagcac gcgccttagc cgccaagtcc tagttcttgc tgccggctct aacgtccgc 60
agtcttcgcc agccagccgt cccgcctgag cgtttgggag gcgtggagcc tgcctgcatg 120
aagtcagcga gagctaagac acccgggaaa cctaccgtga aaaaagggtc ccaaacgaac 180
cttaagacac cagttggggt atactgtagg gtgcgcccac tgggctttcc tgatcaagag 240
tgtttgcata aagtgatcaa taatacaact gttcagcttc atactcctga gggctacaga 300
ctcaaccgaa atggagacta taaggagact cagtattcat ttaaacaagt atttggcact 360
cacaccacc agaagggaact ctttgatgtt gtggctaata ccttgggtcaa tgacctcatt 420
catggcaaaa atgggtcttct ttttacatat ggtgtgacgg gaagtggaaa aactcacaca 480
atgactgggt ctccagggga aggagggtct cttcctcgtt gtttggacat gatctttaac 540
agtatagggt catttcaagc taaacgatat gttttcaaat ctaatgatag gaatagtatg 600
gatatacagt gtgaggttga tgccttatta gaacgtcaga aaagagaagc tatgccaat 660
ccaaagactt cttctagcaa acgacaagta gatccagagt ttgcagatat gataactgta 720

```

caagaattct	gcaaagcaga	agaggttgat	gaagatagtg	tctatgggtgt	atgtgtctct	780
tatattgaaa	tatataataa	ttacatatat	gatctattgg	aagagggtgcc	gtttgatccc	840
ataaaaccca	aacctccaca	atctaaattg	cttcgtgaag	ataagaacca	taacatgtat	900
gttgccagga	gtacagaagt	tgaagtgaag	tctactgagg	aggcttttga	agttttctgg	960
agaggccaga	aaaagagacg	tattgctaag	acccatttga	atcgtgagtc	cagccgttcc	1020
catagcgtgt	tcaacattaa	attagttcag	gctcccttgg	atgcagatgg	agacaatgtc	1080
ttacaggaaa	aagaacaaat	cactataagt	cagttgtcct	tggtagatct	tgctggaagt	1140
gaaagaacta	accggaccag	agcagaaggg	aacagattac	gtgaagctgg	taatattaat	1200
cagtcactaa	tgacgctaag	aacatgtatg	gatgtcctaa	gagagaaaca	aatgtatgga	1260
actaacaaga	tggttccata	tcgagattca	aagttaaccc	atctgtttca	gaactacttt	1320
gatgggggaa	gaaaagtgcg	gatgatcgtg	tgtgtgaacc	ccaaggctga	agattatgaa	1380
gaaaacttgc	aagtcatgag	atgtgcggaa	gtgactcaag	aagttgaaat	agcaagacct	1440
gtagacaagg	caatatgtgg	tttaacgcct	gggaggagat	acagaaacca	gcctcgaggt	1500
ccagttggaa	atgaaccatt	ggttactgac	gtggttttgc	agagttttcc	acctttgccg	1560
tcattgcgaa	ttttggatat	caacgatgag	cagacacttc	caaggctgat	tgaagcctta	1620
gagaaacgac	ataacttacg	acaaatgatg	attgatgagt	ttacaaaaca	atctaattgct	1680
tttaaagctt	tggtacaaga	atgtgacaat	gctgttttaa	gtaaagaaa	ccacatgcaa	1740
gggaaactaa	atgaaaagga	gaagatgatc	tcaggacaga	aattggaaat	agaacgactg	1800
gaaaagaaaa	acaaaacttt	agaatataag	attgagattt	tagagaaaac	aactactatc	1860
tatgagggaag	ataaacgcaa	tttgcaacag	gaacttgaaa	ctcagaacca	gaaacttcag	1920
cgacagtttt	ctgacaaacg	cagattagaa	gccaggttgc	aaggcatggt	gacagaaacg	1980
acaatgaagt	gggagaaaga	atgtgagcgt	agagtggcag	ccaaacagct	ggagatgcag	2040
aataaactct	gggttaaaga	tgaaaagctg	aaacaactga	aggctattgt	tactgaacct	2100
aaaacttgaga	agccagagag	accctctcgg	gagcgagatc	gagaaaaagt	tactcaaaga	2160
tctgtttctc	catcacctgt	gcctttactc	tttcaacctg	atcagaacgc	accaccaatt	2220
cgtctccgac	acagacgatc	acgctctgca	ggagacagat	gggtagatca	taagcccgcc	2280
tctaactatg	aaactgaaac	agtcattgcag	ccacatgtcc	ctcatgccat	cacagtatct	2340
gttgcaaatg	aaaaggcact	agctaagtgt	gagaagtaca	tgctgaccac	ccaggaacta	2400
gcctccgatg	gggagattga	aactaaacta	attaagggtg	atatttataa	aacaaggggt	2460
ggtggacaa	ctgttcagtt	tactgatatt	gagactttaa	agcaagaaat	accaaatggt	2520
agtcgaaaa	gaagatcttc	cacagtagca	cctgcccaac	cagatgggtg	agagtctgaa	2580
tggaaccgat	tagaaacaag	gtgttctgtg	gctgtggaga	tgagagcagg	atcccagctg	2640
ggacctggat	atcagcatca	cgcacaaccc	aagcgcaaaa	agccatgaac	tgacagtccc	2700
agtactgaaa	gaacattttc	atgtgtgtgg	atgatttctc	gaaagccatg	ccagaagcag	2760
tcttccaggt	catcttgtag	aaotccagct	ttgttgaaaa	tcacggacct	cagctacatc	2820
atacactgac	ccagagcaaa	gctttcccta	tggttccaaa	gacaactagt	attcaacaaa	2880
ccttgtatag	tatatgtttt	gccatattta	atattaatag	cagaggaaga	ctcctttttt	2940
catcactgta	tgaatttttt	ataatgtttt	tttaaaatat	atttcatgta	tacttataaa	3000
ctaattcaca	caagtgtttg	tcttagatga	ttaaggaaga	ctatatctag	atcatgtctg	3060
attttttat	gtgacttctc	cagccctggt	ctgaatttct	taaggtttta	taaacaaatg	3120
ctgctattta	ttagctgcaa	gaatgcactt	tagaactatt	tgacaattca	gactttcaaa	3180
ataaagatgt	aaatgactgg	ccaataataa	ccattttagg	aagggtgttt	gaattctgta	3240
tgtatatatt	cactttctga	catttagata	tgccaaaaga	attaaaatca	aaagcactaa	3300
gaaataaaaa	aaaaaaaaaa	aaaa	3324			

<210> 160

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 160

caaagctttc cctatgggtc aaagacaact agtattcaac aaaccttgta tagtgtatgt 60

<210> 161

<211> 1536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 161

acagagcttc	aaaaaaagag	cgggacaggg	acaagcgtat	ctaagaggct	gaacatgaat	60
ccacagatca	gaaatccgat	ggagcggatg	tatcgagaca	cattctacga	caactttgaa	120
aacgaaccca	t cctctatgg	tcggagctac	acttggctgt	gctatgaagt	gaaaataaag	180
agggggccgt	caaatctcct	ttgggacaca	ggggctcttc	gaggccaggt	gtatttcaag	240
cctcagtacc	acgcagaaat	gtgcttcctc	tcttggttct	gtggcaacca	gctgcctgct	300
tacaagtgtt	tcagatcac	ctggtttgta	tcctggaccc	cctgcccgga	ctgtgtggcg	360
aagctggccg	aattcctgtc	tgagcacccc	aatgtcacc	tgaccatctc	tgccgcccgc	420
ctctactact	actgggaaag	agattaccga	agggcgctct	gcaggctgag	tcaggcagga	480
gcccgcgtga	c gatcatgga	ctatgaagaa	tttgcatact	gctgggaaaa	ctttgtgtac	540
aatgaaggtc	agcaattcat	gccttggtac	aaattcgatg	aaaattatgc	attcctgcac	600
cgcacgctaa	aggagattct	cagataacctg	atggatccag	acacattcac	tttcaacttt	660
aataatgacc	ctttggctct	tcgacggcgc	cagacctact	tgtgctatga	ggtggagcgc	720
ctggacaatg	gcacctgggt	cctgatggag	cagcacatgg	gctttctatg	caacgaggct	780
aagaatcttc	tctgtggctt	ttacggccgc	catgcggagc	tgcgcttctt	ggacctgggt	840
ccttctttgc	agttggaccc	ggcccagatc	tacagggtca	cttggttcat	ctcctggagc	900
ccctgcttct	cttggggctg	tgccggggaa	gtgcgtgcgt	tccttcagga	gaacacacac	960
gtgagactgc	gcctcttcgc	tgcccgcctc	tatgattacg	acccctata	taaggaggcg	1020
ctgcaaagtc	tgcgggatgc	tggggcccaa	gtctccatca	tgacctacga	tgagtttgag	1080
tactgctggg	acacctttgt	gtaccgccag	ggatgtccct	tccagccctg	ggatggacta	1140
gaggagcaca	gccaaagcct	gagtgggagg	ctgcggggca	ttctccagaa	tcagggaaac	1200
tgaaggatgg	gcctcagtc	ctaaggaagg	cagagacctg	ggttgagcag	cagaataaaa	1260
gatcttcttc	caagaaatgc	aaacagaccg	ttcaccacca	tctccagctg	ctcacagaca	1320
ccagcaaagc	aatgtgctcc	tgatcaagta	gatttttttaa	aaatcagagt	caattaattt	1380
taattgaaaa	tttctcttat	gttccaagtg	tacaagagta	agattatgct	caatattccc	1440
agaatagttt	tcaatgtatt	aatgaagtga	ttaattgggt	ccatatttag	actaataaaa	1500
cattaagaat	cttccataat	tgtttccaca	aacact	1536		

<210> 162

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 162

tgctcacaga	caccagcaaa	gcaatgtgct	cctgatcaag	tagatttttt	aaaaatcaga	60
------------	------------	------------	------------	------------	------------	----

<210> 163

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 163

cgtagagttc	ggccgaagga	acctgaccca	ggctctgtga	ggaggcaagg	ttttcagggg	60
acaggccaac	ccagaggaca	ggattccctg	gaggccacag	aggagcacca	aggagaagat	120
ctgcctgtgg	gtcttcattg	cccagctcct	gccacactc	ctgcctgctg	ccctgacgag	180
agtcatcatg	tctcttgagc	agaggagtct	gcactgcaag	cctgaggaa	cccttgaggc	240
ccaacaagag	gccctggggc	tggtgtgtgt	gcaggctgcc	gcctcctct	cctctcctct	300
ggtcctgggc	accctggagg	agggtgccac	tgctgggtca	acagatcctc	cccagagtcc	360
tcaggggagc	tcgccttttc	ccactaccat	caacttcact	cgacagaggc	aaccagtgga	420
gggttccagc	agccgtgaag	aggagggggc	aagcacctct	tgtatcctgg	agtccttggt	480
ccgagcagta	atcactaaga	agggtggctga	tttggttggg	tttctgctcc	tcaaataatcg	540
agccagggag	ccagtcacaa	aggcagaaat	gctggagagt	gtcatcaaaa	attacaagca	600

ctgttttccct	gagatcttctg	gcaaagcctc	tgagtccttg	cagctgggtct	ttggcattga	660
cgtgaaggaa	gcagacccca	ccggccactc	ctatgtcctt	gtcacctgcc	taggtctctc	720
ctatgatggc	ctgctgggtg	ataatcagat	catgcccaag	acaggcttcc	tgataattgt	780
cctgggtcatg	attgcaatgg	agggcggcca	tgctcctgag	gaggaaa tct	gggaggagct	840
gagtgatgatg	gaggtgtatg	atgggagggg	gcacagtgcc	tatggggagc	ccaggaagct	900
gctcacccaa	gatttgggtgc	aggaaaagta	cctggagtag	cggcagggtgc	cggacagtga	960
tcccgcacgc	tatgagttcc	tgtgggggtcc	aagggccctt	gctgaaa cca	gctatgtgaa	1020
agtccttgag	tatgtgatca	aggtcagtgc	aagagttcgc	tttttct tcc	catccctgcg	1080
tgaagcagct	ttgagagagg	aggaagaggg	agtctgagca	tgagttgcag	ccagggccag	1140
tgggagggggg	actgggccag	tgcaccttcc	agggccgcgt	ccagcagctt	cccctgcctc	1200
gtgtgacatg	aggcccatte	ttcactctga	agagagcggt	cagtgtt ctc	agtagtaggt	1260
ttctgttcta	tgggtgact	tggagattta	tctttgttct	cttttggaat	tgttcaaagt	1320

ttttttttta	aggggatggtt	gaatgaactt	cagcatccaa	gtttatgaat	gacagcagtc	1380
acacagttct	gtgtatatag	tttaagggtg	agagtcctgt	gttttat tca	gattgggaaa	1440
tccattctat	tttgtgaatt	gggataataa	cagcagtggg	ataagta ctt	agaaatgtga	1500
aaaatgagca	gtaaaataga	tgagataaag	aactaaagaa	attaaga gat	agtcaattct	1560
tgccttatac	ctcagtcctat	tctgtaaaaa	ttttaaagat	atatgca tac	ctggatttcc	1620
ttggcttctt	tgagaatgta	agagaaatta	aatctgaata	aagaatt ctt	cctgttaaaa	1680
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aa 172 2		

<210> 164

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 164

cagattggga	aatccattct	attttgtgaa	ttgggataat	aacagcagtg	gaataagtag	60
------------	------------	------------	------------	------------	------------	----

<210> 165

<211> 2334

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004994

<400> 165

agacacctct	gccctcacca	tgagcctctg	gcagccccctg	gtcctgggtgc	tcctgggtgct	60
gggctgctgc	t ttgctgccc	ccagacagcg	ccagtccacc	cttgtgc tct	tccttgagga	120
cctgagaacc	aatctcaccg	acaggcagct	ggcagaggaa	tacctgt acc	gctatgggtta	180
cactcgggtg	gcagagatgc	gtggagagtc	gaaatctctg	gggectgcgc	tgctgcttct	240
ccagaagcaa	ctgtccctgc	ccgagaccgg	tgagctggat	agcgcca cgc	tgaaggccat	300
gcgaacccca	cgggtgcgggg	tcacagacct	gggcagattc	caaacct ttg	agggcgacct	360
caagtggcac	caccacaaca	tcacctattg	gatccaaaac	tactcggaag	acttgccgcg	420
ggcggtgatt	gacgacgcct	ttgcccgcgc	cttcgcactg	tggagcgcgg	tgacgcgcgt	480
caccttcact	cgcgtgtaca	gccgggaacg	agacatcgtc	atccagt ttg	gtgtcgcgga	540
gcacggagac	gggtatccct	tcgacgggaa	ggacgggctc	ctggcac acg	cctttccctcc	600
tggccccggc	attcaggagg	acgcccattt	cgacgatgac	gagttgt ggt	ccctgggcaa	660
gggcgtcggtg	gttccaactc	ggtttggaat	cgcagatggc	gcggcct gcc	acttcccctt	720
catcttcgag	ggccgctcct	actctgcctg	caccaccgac	ggtcgct cgc	acggcttgcc	780
ctggtgcagt	accacggcca	actacgacac	cgacgaccgg	tttggct tct	gccccagcga	840
gagactctac	accggggacg	gcaatgctga	tgggaaaccc	tgccagt ttc	cattcatctt	900
ccaaggccaa	t cctactccg	cctgcaccac	ggacggtcgc	tccgacggct	accgctgggtg	960
cgccaccacc	gccaaactacg	accgggacaa	gctcttcggc	ttctgcc cga	cccagagctga	1020
ctcgacgggtg	atgggggggca	actcggcggg	ggagctgtgc	gtcttcc cct	tcactttcct	1080
gggtaaggag	tactcgacct	gtaccagcga	gggccgcgga	gatgggc gcc	tctgggtgcgc	1140
taccacctcg	aactttgaca	gcgacaagaa	gtggggcttc	tgcccggacc	aaggatacag	1200

```

tttgttctct  gtggcgggcg  atgagttcgg  ccacgcgctg  ggcttagatc  attcctcagt  1260
gccggaggcg  ctcatgtacc  ctatgtaccg  cttoactgag  gggcccCct  tgcataagga  1320
cgacgtgaat  ggcattccgg  acctctatgg  tcttcgcctt  gaacctgagc  cacggcctcc  1380
aaccaccacc  acaccgcagc  ccacggctcc  cccgacgggt  tgccccaccg  gacccccac  1440
tgtccacccc  tcagagcgcc  ccacagctgg  cccacaggt  ccccccctcag  ctggccccac  1500
aggtcccccc  actgctggcc  cttctacggc  cactactgtg  cctttgagtc  cgggtggacga  1560
tgcttgcaac  gtgaacatct  tcgacgccat  cgcggagatt  ggggaacCagc  tgtatttggt  1620
caaggatggg  aagtactggc  gattctctga  gggcaggggg  agccggCcg  agggccctt  1680
ccttatcgcc  gacaagtggc  ccgcgctgcc  ccgcaagctg  gactcgggtct  ttgaggagcc  1740
gctctccaag  aagcttttct  tcttctctgg  gcgccaggtg  tgggtgtaca  caggcgcgtc  1800
gggtgctggg  ccgaggcgtc  tggacaagct  gggcctggga  gccgacgtgg  cccaggtgac  1860
cggggccctc  cggagtggca  gggggaagat  gctgctgttc  agcgggCggc  gcctctggag  1920

gttcgacgtg  aaggcgcgca  tgggtggatcc  ccggagcgcc  agcgaggtgg  accggatggt  1980
ccccgggggt  cctttggaca  cgcacgacgt  cttccagtac  cgagagaag  cctatttctg  2040
ccaggaccgc  ttctactggc  gcgtgagttc  ccggagttag  ttgaaccagg  tggaccaagt  2100
gggctacgtg  acctatgaca  tcttcgagtg  ccctgaggac  tagggctccc  gtctgtcttt  2160
gcagtgccat  gtaaattccc  actgggacca  acctggggga  aggagccagt  ttgccggata  2220
caaactggta  ttctgttctg  gaggaaaggg  aggagtggag  gtgggctggg  cctctcttcc  2280
tcacctttgt  tttttgttgg  agtgtttcta  ataaacttgg  attctctaac  cttt  2334

```

<210> 166

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004994

<400> 166

```

ggccctctct  tctcaccttt  gttttttggt  ggagtgtttc  taataaaCtt  ggattctcta  60

```

<210> 167

<211> 5329

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base

<222> 1 ... 5329

<223> n = a,c,g, or t

<300>

<308> NM_005063

<400> 167

```

gtggtgtcgg  tgtcggcagc  atccccggcg  ccctgctggc  gtcgccggag  ccctcggcct  60
ctgttctcct  cccctcccg  cccttacctc  cagcggggac  cgcccgcgcc  agtcaactcc  120
tcgcactttg  ccctgtcttg  gcagcggata  aaagggggct  gaggaaatac  cggacacgtc  180
caccggttgc  cagctctagc  ctttaaattc  ccggctcggg  acctccacgc  accgggctag  240
cgccgacaac  cagctagcgt  gcaaggcgcc  gcggctcagc  gcgtaccggc  gggcttcgaa  300
accgcagtcc  tccggcgacc  ccgaactccg  ctccggagcc  tcagcccCct  ggaaagtgat  360
cccggcatcg  gagagccaag  atgccggccc  acttgtctga  ggacgataatc  tctagctcct  420
ataccaccac  caccaccatt  acagcgcttc  cctccagggt  cctgcagaat  ggaggagata  480
agttggagac  gatgcccctc  tacttggaag  acgacattcg  cctgataata  aaagatgata  540
tatatgaccc  cactacaag  gataaggaag  gcccaagccc  caaggttgaa  tatgtctgga  600
gaaacatcat  ccttatgtct  ctgtacact  tgggagccct  gtatgggatac  actttgatcc  660
ctacctgcaa  gttctacacc  tggctttggg  gggatttcta  ctattttgtc  agtgccctgg  720
gcataacagc  aggagctcat  cgtctgtgga  gccaccgctc  ttacaaagct  cggctgcccc  780
tacggctctt  tctgatcatt  gccaacacaa  tggcattcca  gaatgatgtc  tatgaatggg  840
ctcgtgacca  ccgtgcccac  cacaagtttt  cagaaacaca  tgctgatcct  cataattccc  900

```

gacgtggcctt	tttctttctct	cacgtggggtt	ggctgcttgt	gcgcaaacac	ccagctgtca	960
aagagaaggg	gagtacgcta	gacttgtctg	acctagaagc	tgagaaactg	gtgatgttcc	1020
agaggaggta	ctacaaacct	ggcttgcctg	tgatgtgctt	catcctgccc	acgcttgtgc	1080
cctggtatatt	ctgggggtgaa	acttttcaaa	acagtgtgtt	cgttgcact	ttcttgcgat	1140
atgctgtgggt	gcttaatgcc	acctggctgg	tgaacagtgc	tgcccaactc	ttcggatata	1200
gtccttatga	caagaacatt	agccccggg	agaatatcct	ggtttcaact	ggagctgtgg	1260
gtgagggcctt	ccacaactac	caccactcct	ttccctatga	ctactctgcc	agttagtacc	1320
gctggccacat	caacttcacc	acattcttca	ttgattgcat	ggccgcctc	ggctgggcct	1380
atgaccggaa	gaaagtctcc	aaggccgcca	tcttggccag	gattaaaaga	accggagatg	1440
gaaactacaa	gagtggctga	gtttgggggtc	cctcagggttc	ctttttcaaa	aaccagccag	1500
gcagagggttt	taatgtctgt	ttattaacta	ctgaataatg	ctaccaggat	gctaaagatg	1560
atgatgttaa	cccattccag	tacagtattc	ttttaaaatt	caaaagtatt	gaaagccaac	1620
aactctgcct	ttatgatgct	aagctgatata	tatttcttct	cttatcctct	ctctcttcta	1680
ggcccattgt	cctccttttc	actttaatcg	ccctcctttc	ccttattgcc	tcccaggcaa	1740
gcagctgggtc	agtcttttgc	cagtgtccag	cttccaaagc	ctagacaacc	tttctgtagc	1800
ctaaaacgaa	tggtctttgc	tccagataac	tctctttcct	tgagctgttg	tgagctttga	1860
agtagggtgc	ttgagctaga	gataaaacag	aatcttctgg	gtagtcctct	gttgattatc	1920
ttcagcccag	gcttttgcga	gatggaatgg	aaaagcaact	tcatttgaca	caaagcttct	1980
aaagcnagggt	aaattgtcgg	gggagagagt	tagcatgtat	gaatgtaagg	atgagggaag	2040
cgaagggaacc	tctcgccatg	atcagacata	cagctgccta	cctaattgagg	acttcaagcc	2100
ccaccacata	gcatgcttcc	tttctctcct	ggctcggggt	aaaaagtggc	tgcggtgttt	2160
ggcaatgcta	attcaatgcc	gcaacatata	gttgaggccg	aggataaaga	aaagacattt	2220
taagtttgta	gtaaaagtgg	tctctgctgg	ggaagggttt	tcttttcttt	ttttctttaa	2280
taacaaggag	atttcttagt	tcataatatca	agaagtcttg	aagttgggtg	tttccagaat	2340
tggtaaaaac	agcagctcat	agaattttga	gtattccatg	agctgctcat	tacagttctt	2400
tctcttttct	gctctgccat	cttcaggata	ttggttcttc	ccctcatagt	aataagatgg	2460
ctgtggcatt	tccaaacatc	caaaaaagg	gaaggattta	aggagggtgaa	gtcgggtcaa	2520
aaataaaaata	tatatacata	tatacattgc	ttagaacgtt	aaactattag	agtatttccc	2580
ttccaaagag	ggatgtttgg	aaaaaactct	gaaggagagg	agggaattagt	tgggatgcca	2640
atttctctct	cactgctgga	catgagatgg	agaggctgag	ggacaggatc	tataggcagc	2700
ttctaagagc	gaacttcaca	taggaaggga	tctgagaaca	cgttcagggg	ttgagaagggt	2760
tactgagtga	gttattggga	gtcttaataa	actagatatt	aggccaattc	attaattagt	2820
tccagtttct	ccttgaaatg	agtaaaaact	agaaggcttc	tctccacagt	gttgtgcccc	2880
ttcactcatt	tttttttgag	gagaaggggg	tctctgttaa	catctagcct	aaagtataca	2940
aactgcctgg	ggggcagggt	taggaatctc	ttcactacc	tgattcttga	ttcctggctc	3000
taccctgtct	ttgccttttc	tttgaccaga	tctttctctt	ccctgaa.cgt	tttcttcttt	3060
ccctggacag	gcagcctcct	ttgtgtgtat	tcagaggcag	tgatgacttg	ctgtccaggc	3120
agctccctcc	tgcacacaga	atgctcaggg	tcactgaacc	actgcttctc	ttttgaaagt	3180
agagctagct	gccactttca	cgtggcctcc	gcagtgtctc	cacctacacc	cctgtgctcc	3240
cctgccacac	tgatggctca	agacaaggct	ggcaaacctc	cccagaaaca	tctctggccc	3300
agaaagcctc	tctctccctc	cctctctcat	gagaagccaa	gcgctcatgt	tgagccagtg	3360
ggccagccac	agagcaaaaag	agggtttatt	ttcagtcctc	tctctctggg	tcagaaccag	3420
agggcatgct	gaatgcctcc	tgcttacttg	gtgagggtgc	ccgcctgag	tcagtgtctt	3480
cagctggcag	tgcaatgctt	gtagaagttag	gaggaaacag	ttctcac.tgg	gaagaagcaa	3540
gggcaagaac	ccaagtgcct	cacctcgaaa	ggaggccctg	ttccctggag	tcagggtgaa	3600
ctgcaaagct	ttggctgaga	cctgggattt	gagataccac	aaacctgct	gaacacagtg	3660
tctgttcagc	aaactaacca	gcattcccta	cagcctaggg	cagacaa.tag	tatagaagtc	3720
tggaaaaaaa	caaaaacaga	atttgagaac	cctggaccac	tctgtctcct	gtagctcagt	3780
catcaaagca	gaagtctggc	tttgctctat	taagattgga	aatgtacact	accaaact	3840
cagtccactg	ttgagcccca	gtgctggaag	ggaggaaggc	ctttctt.ctg	gtttaattgc	3900
gtagaggcta	caggggttag	cctggactaa	aggcatcctt	gtctttgagc	tattcacctc	3960
agtagaaaag	gatctaaggg	aagatcactg	tagtttagtt	ctgttgac.cct	gtgcacctac	4020
cccttggaag	tgtctgctgg	tatttctaata	tccacaggtc	atcagatgcc	tgcttgataa	4080
tatataaaca	ataaaaaaaa	ctttcacttc	ttcctattgt	aatcgtgtgc	catggatctg	4140
atctgtacca	tgaccctaca	taaggctgga	tggcacctca	ggctgagggc	cccaatgtat	4200
gtgtggctgt	gggtgtgggt	gggagtgtgt	ctgctgagta	aggaaca.cga	ttttcaagat	4260
tctaaagctc	aattcaagtg	acacattaat	gataaactca	gatctga.tca	agagtccgga	4320
tttctaacag	tccttgcttt	gggggggtgt	ctggcaactt	agctcaggtg	ccttacatct	4380
tttctaataca	cagtgttgca	tatgagcctg	ccctcactcc	ctctgcagaa	tccctttgca	4440
cctgagaccc	tactgaagtg	gctggtagaa	aaaggggcct	gagtggaggga	ttatcagtat	4500
cacgatttgc	aggattccct	tctgggcttc	attctggaaa	cttttgt.tag	ggctgctttt	4560

```

cttaagtgcc cacatttgat ggagggtgga aataatttga atgtatttga tttataagtt 4620
tttttttttt tttgggttaa aagatggttg tagcatttaa aatggaaaat tttctccttg 4680
gtttgctaagt atcttgggtg tattctctgt aagtgtagct caaatagggtc atcatgaaag 4740
gttaaaaaag cgagggtggcc atgttatgct ggtggttgcc agggcctcca accactgtgc 4800
cactgacttg ctgtgtgacc ctgggcaagt cacttaacta taagggtgct cagttttcct 4860
tctgttaaaa tggggataat aatactgacc tacctcaaag ggcagttt tg aggcattgact 4920
aatgcttttt agaaagcatt ttgggatcct tcagcacagg aattctcaag acctgagtat 4980
tttttataat aggaatgtcc accatgaact tgatacgtcc gtgtgtcc ca gatgctgtca 5040
ttagtctata tggttctcca agaaactgaa tgaatccatt ggagaagcgg tggataacta 5100
gccagaca aa atttgagaat acataaacia cgcattgcc cggaaaca ta cagaggatgc 5160
cttttctgtg attgggtggg attttttccc tttttatgtg ggatatagta gttacttgtg 5220
acaagaataa ttttggaata atttctatta atatcaactc tgaagcta at tgtactaatc 5280
tgagattgtg tttgttcata ataaaagtga agtgaatctg attgcactg 5329

```

<210> 168

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005063

<400> 168

```

aataatgcta ccaggatgct aaagatgatg atgttaaccc attccagtac agtattcttt 60

```

<210> 169

<211> 634

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 169

```

cggctgagag gcagcgaact catctttgcc agtacaggag cttgtgccgt ggcccacagc 60
ccacagccca cagccatggg ctgggacctg acggtgaaga tgctggcggg caacgaattc 120
cagggtgtccc tgagcagctc catgtcgggtg tcagagctga aggcgcagat caccagaag 180
attggcgtgc acgccttcca gcagcgtctg gctgtccacc cgagcgggtgt ggcgctgcag 240
gacagggtcc ccttgccag ccagggcctg ggccctggca gcacggctct gctgggtggtg 300
gacaaatgcg acgaacctct gagcatcctg gtgaggaata acaagggcgg cagcagcacc 360
tacgaggtcc ggctgacgca gaccgtggcc cacctgaagc agcaagtga ggggctggag 420
ggtgtgcagg acgacctgtt ctggctgacc ttcgagggga agcccctgga ggaccagctc 480
ccgctggggg agtacggcct caagcccctg agcaccgtgt tcatgaatct gcgcctgcgg 540
ggaggcggca cagagcctgg cgggcggagc taagggcctc caccagcatc cgagcaggat 600
caagggccgg aaataaaggc tgttgtaaga gaat 634

```

<210> 170

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 170

```

tggtggtgga caaatgcgac gaacctctga gcatcctggt gaggaataac aagggccgca 60

```

<210> 171

<211> 1339

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 171

```

gaattccgat tagtgtgatc tcagctcaag gcaaagggtgg gata tcatgg catctatctg 60
ggttggacac cgaggaacag taagagatta tccagacttt agcc catcag tggatgctga 120
agctattcag aaagcaatca gaggaattgg aactgatgag aaaa tgctca tcagcattct 180
gactgagagg tcaaatgcac agcggcagct gattgttaag gaat atcaag cagcatatgg 240
aaagggagctg aaagatgact tgaagggtga tctctctggc cact ttgagc atctcatggg 300
ggccctagtg actccaccag cagtctttga tgcaaagcag ctaa agaaat ccatgaaggg 360
cgcggaaca aacgaagatg ccttgattga aatcttaact acca ggacaa gcaggcaaat 420
gaaggatata tctcaagcct attatacagt atacaagaag agtc ttggag atgacattag 480
ttccgaaaca tctggtgact tccggaaagc tctgttgact ttgg cagatg gcagaagaga 540
tgaaagtctg aaagtggatg agcatctggc caaacaagat gccc agattc tctataaagc 600
tggtgagaac agatggggca cggatgaaga caaattcact gaga tcctgt gtttaaggag 660
ctttcctcaa ttaaaactaa catttgatga atacagaaat atca gccaaa aggacattgt 720
ggacagcata aaaggagaat tatctgggca ttttgaagac ttac tgttgg ccatagttaa 780
ttgtgtgagg aacacgcggc cttttttagc cgaaagactg catc gagcct tgaagggtat 840
tggaactgat gagtttactc tgaaccgaat aatgggtgtcc agat cagaaa ttgacctttt 900
ggacattcga acagagttca agaagcatta tggctattcc ctat attcag caattaaatc 960
ggatacttct ggagactatg aaatcacact cttaaaaatc tgtgtgtggag atgactgaac 1020
caagraagata atctccaaag gtccacgatg ggctttccca acag ctccac cttacttctt 1080
ctcactactat ttaagagaac aagcaaatat aaacagcaac ttgt gtccct aacaggaatt 1140
ttcattgttc tataacaaca acaacaaaag cgattattat tttagagcat ctcatattata 1200
atgtagcagc tcataaatga aattgaaaat ggtattaaag atct gcaact actatccaac 1260
ttatatttct gctttcaaag ttaagaatct ttatagtctt actc cattaa atataaagca 1320
agataataaa acggaattc 1339

```

<210> 172

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 172

```

ttcagcaatt aaatcggata cttctggaga ctatgaaatc acac tcttaa aaatctgtgg 60

```

<210> 173

<211> 1582

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 173

```

ccgagctgtg cttgtggctg cggctgctaa ctggctgcgc acagggagct gtcaccatgc 60
ctcactcgta cccagccctt tctgctgagc agaagaagga gttgtctgac attgccctgc 120
ggatgttagc cccgggcaaa ggcattctgg ctgcggatga gtct gtaggc agcatggcca 180
agcggctgag ccaaattggg gtggaaaaca cagaggagaa ccgc cggctg taccgccagg 240
tcctgttcag tgctgatgac cgtgtgaaaa agtgcatggg aggcgtcatt ttcttccatg 300
agacccctcta ccagaaagat gataatgggtg ttcccttcgt ccga accatc caggataagg 360
gcatcgtcgt gggcatcaag gttgacaagg gtgtggtgcc tcta gctggg actgatggag 420
aaacccaccac tcaagggtcg gatgggctct cagaacgctg tgcc caatac aagaaggatg 480
gtgctgactt tgccaagtgg cgctgtgtgc tgaaaatcag tgag cgtaca cctctgcac 540
ttgccattct ggagaacgcc aacgtgctgg cccgttatgc cagt atctgc cagcagaatg 600
gcattgtgcc tattgtggaa cctgaaatat tgctgatgg agac cagcac ctcaaactgt 660
gtcagtatgt tacagagaag gtcttggctg ctgtgtacaa ggcc ctgagt gaccatcatg 720

```



```

tatacCtggga ggggaccctg ctcaagccca acatggtgac Cccgggccaat gcctgtccca 780
tcaagtatac cccagaggag attgccatgg caactgtcac Tgccctgcgt cgactgtgc 840
ccccagctgt cccaggagtg accttcctgt ctgggggtca gagcgaagaa gaggcatcat 900
tcaacCtcaa tgccatcaac cgetgcccc ttccccgacc Ctgggcgctt accttctcct 960
atgggcgtgc cctgcaagcc tctgcaactca atgcctggcg agggcaacgg gacaatgctg 1020
gggctgccac tgaggagtgc atcaagcggg ctgagggtgaa tgggcttgca gccagggca 1080
agtatgaagg cagtggagaa gatggtggag cagcagcaca gtcactctac attgccaacc 1140
atgcctactg agtatccact ccataccaca gcccttgcc cagccatctg caccacttt 1200
tgcttgtagt catggccagg gccaaatagc tatgcagagc agagatgcct tcacctggca 1260
ccaacttgtc ttcttttctc tcttcccttc cctctctca ttgctgcacc tgggaccata 1320
ggatgggagg ataggagacc cctcatgact gagggcagaa gaaattgcta gaagtcagaa 1380
caggatggct gggctctccc ctacctcttc cagctccac aattttccca tgatgagga 1440
gcttctccct gggctctcct tcttgctgc cctgtctcct gggatcagag ggtagtacag 1500
aagccctgac tcatgccttg agtacatacc atacagcaaa taaatggtag caaaacaaaa 1560
aaaaaAAAAA aaaaaaaaaa aa 1582

```

<210> 174

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 174

```

gagggtagta cagaagccct gactcatgcc ttgagtacat accatacagc aaataaatgg 60

```

<210> 175

<211> 451

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 175

```

acttcCctgt tcactttggt tccagcatcc tgtccagcaa agaagcaatc agccaaaatg 60
atacctggag gcttatctga ggccaaaccc gccactccag aaatccagga gattgttgat 120
aagggtaaac cacagcttga agaaaaaaca aatgagactt atggaaaatt ggaagctgtg 180
cagtataaaa ctcaagttgt tgctggaaca aattactaca ttaaggtagc agcagggtgat 240
aataaatata tgcacttgaa agtattcaaa agtcttccc gacaaaatga ggacttggtg 300
cttactggat accagggtga caaaaacaag gatgacgagc tgacgggctt ttagcagcat 360
gtaccCaaag tgttctgatt ctttcaactg gctactgagt catgatcctt gctgataaat 420
ataaccatca ataaagaagc attcttttcc a 451

```

<210> 176

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005213

<400> 176

```

aactgggctac tgagtcatga tccttgctga taaatataac catcaataaa gaagcattct 60

```

<210> 177

<211> 366

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 177

```

gtcagctcag cctccaaagg agccagcctc tccccagttc ctgaaatcct gagtggtgcc 60
tgccagtcgc catgagaact tcctaccttc tgctgtttac tctctgctta cttttgtctg 120
agatggcctc aggtggtaac tttctcacag gccttggcca cagatctgat cattacaatt 180
gcgtcagcag tggagggcaa tgtctctatt ctgcttccc gatctttacc aaaattcaag 240
gcacctgtta cagaggggaag gccaaagtgt gcaagtgcgc tgggagtgac cagaagaaat 300
gacgcagaag tgaaatgaac tttttataag cattctttta ataaaggaaa attgcttttg 360
aagtat 366

```

<210> 178

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 178

```

gggagtgacc agaagaaatg acgcagaagt gaaatgaact ttttataagc attcttttaa 60

```

<210> 179

<211> 1519

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005326

<400> 179

```

ctgcctcgga acgctgtccc ccgcagcgac ggcccgttcc acctcgcgat ctgccgggta 60
cccgggcggc gtggcgctcg gcctccaggg atccactgtg cgggtgccaa aaagaggcgg 120
aggctcgcgg cacagctctc ccggcgcgac tctcgggccc ccgccgcgcg tcccaggccc 180
gtctcccggc ccgtggcagt cggggctcgc ggacaaaaca agttgagcgc gagcgcgttg 240
attggttggc ggacggtgcg aggtggacgc tgattggctg agggcagcgc gaggcgggcg 300
ctgattggct gcgacgcgcc gacgccggtg ttttgcagtc ctgggcagct cggcagtcga 360
gcccggcccc ggtcatggtg gtgggcccag ggctgctcgg ccgccgcagc ctgcgccgcg 420
tgggagccgc ctgcgccgcg cgaggccctc gtccagccct gctgggagtt ttctgccaca 480
cagattttcg gaagaacctg acctgggacg agggcaccat gaaggtagag gtgctgcctg 540
ccctgaccga caactacatg tacctggtca ttgatgatga gaccaaggag gctgccattg 600
tggatccggt gcagccccag aaggtcgtgg acgcggcgag aaagcacggg gtgaaactga 660
ccacagtgtc caccaccac caccactggg accatgctgg cggaatgag aaactggtca 720
agctgggagtc gggactgaag gtgtacgggg gtgacgaccg tatcggggcc ctgactcaca 780
agatcactca cctgtccaca ctgcaggtgg ggtctctgaa cgtcaagtgc ctggcgaccc 840
cgtgccacac ttcaggacac atttgttact tcgtgagcaa gcccgaggc tcggagcccc 900
ctgccgtgtt cacaggtgac acctgttttg tggctggctg cgggaagtcc tatgaaggga 960
ctgcggatga gatgtgtaaa gctctgctgg aggtcttggg ccggctcccc ccggacacaa 1020
gagtctactg tggccacgag tacaccatca acaacctcaa gtttgacgc cacgtggagc 1080
ccggcaatgc cgccatccgg gagaagctgg cctgggcca ggagaagtac agcatcgggg 1140
agcccacagt gccatccacc ctggcagagg agtttaocta caacccttc atgagagtga 1200
gggagaagac ggtgcagcag caccaggtg agacggacc ggtgaccacc atgcggggccg 1260
tgccgaggga gaaggaccag ttcaagatgc cccgggactg aggcgcgcc ctgaccttcag 1320
cggatttggg gattaggctc ttttaggtaa ctggcttccc tgctgggtccg tcggggaaat 1380
tcagtcttga tttaacctta attttacagc ccttggcttg tgttatcgga cattctaatt 1440
catatttata agagaagttt aacaagtatt tattcccata aaaaaaaaaa aaaaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa 1519

```

<210> 180

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005326

<400> 180

cttgtgttat cggacattct aatgcataatt tataagagaa gtttaacaag tattttattcc 60

<210> 181

<211> 3378

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005461

<400> 181

acagctgcac	cgccgagctg	cgagcggcctg	cgagcgcagag	agcgttaagag	caagagagct	60
agagagcgag	caacggggcac	tgcceccacg	cctcccctca	gccccaccgc	gcgctccgct	12 0
tgcctctcca	ccccgcccga	ctctacccgg	cccggtccct	gcgcggggcac	agcccagagc	18 0
tctggggcgg	tgcaggcagc	ctcgggactc	tccggcgcg	cgccgcgtcc	ccagacaaag	24 0
gcttggccgg	cggccccggc	ccgctgcgcc	ctcgtccccc	gcctccccag	ctcttctccg	30 0
ctcttccccc	ccgcgcttgg	ctcggcgcg	tccggccggc	cgcaaagttt	cccgggcggc	36 0
agcggcggt	gcgcctcgct	tcagcgatgg	ccgcggagct	gagcatgggg	ccagagctgc	42 0
ccaccagccc	gctggccatg	gagtatgtca	acgacttcga	cctgctcaag	ttcgacgtga	48 0
agaaggagcc	actggggcgc	gcggagcgct	cgggcaggcc	ctgcacacgc	ctgcagccag	54 0
ccggctcgg	gtcctccaca	ccgctcagca	ctccgtgtag	ctccgtgccc	tcgtcgccca	60 0
gcttcagccc	gaccgaacag	aagacacaac	tcgaggatct	gtactggatg	gcgagcaact	66 0
accagcagat	gaaccccgag	gcgctcaaac	tgacgcccga	ggacgcgggtg	gaagcgctca	72 0
tcggctcgca	cccagtgcca	cagccgctgc	aaagcttcga	cagctttcgc	ggcgctcacc	78 0
accaccacca	tcaccaccac	cctcaccggc	accacgcgta	ccggggcgcc	ggcgtaggcc	84 0
acgacgagct	gggcccgcac	gctcaccggc	accatcacca	tcataccaa	gcgtcgccgc	90 0
cgccgtccag	cgccgctagc	ccggcgcaac	agctgcccac	tagccacccc	gggcccgggc	96 0
cgcacgcgac	ggcctcggcg	acggcgccgg	gcggcaacgg	cagcgtggag	gaccgcttct	10 20
ccgacgacca	gctcgtgtcc	atgtccgtgc	gcgagctgaa	ccgccacctg	cggggcttca	10 80
ccaaggacga	ggtgatccgc	ctgaagcaga	agcggcggac	cctgaagaac	cggggctacg	11 40
cccagtcctg	caggtataaa	cgcgtccagc	agaagcacca	cctggagaat	gagaagacgc	12 00
agctcattca	gcaggtggag	cagcttaagc	aggaggtgtc	ccggctggcc	cgcgagagag	12 60
acgcctacaa	ggtcaagtgc	gagaaactcg	ccaactccgg	cttcagggag	gcgggctcca	13 20
ccagcgacag	cccctcctct	cccagattct	ttctgtgagt	cgtggccgggt	cctggccccc	13 80
gcccttgccc	cggcccggac	tccctgtccc	acgtccctag	tcccagacta	ccccggacct	14 40
tgtccctgcc	gcggcccag	ccttgacctg	tttgacttga	gcgagagggg	ggaaggcgcc	15 00
gcgggcccgc	ggcgacgggc	gggtgcgcgg	gcgggcaggg	gaccttggt	aaggcgagag	15 60
tagcgcacgc	cagcgccgcc	tcctagactc	gagcagagcc	ggagagagag	acgagagggt	16 20
gggagggtccc	ggagtaactt	ctctccaggc	tgaaggggcg	cgaggcatag	tcccagagaag	16 80
tcaccaaggc	catctggaga	ctcctggctt	tctgaacttt	gcgcgttaag	ccgggacagc	17 40
tgttttgctg	cccggagagt	agtccgcgcc	aggaagagag	caacgaggaa	aggagaggga	18 00
ctctggcgct	ccggcaggcg	agaggcgagg	ctgagcgaaa	gaaggaagga	cagacggacc	18 60
tgtctgtcag	agttcggaga	acactggctc	tcagccctga	gacacaggcc	tcagtttagga	19 20
cgtctggcgc	ccaaatctca	tcagttttat	tgctgtctcg	attatataga	aaaatacaaa	19 80
aaatctgcat	taaaaatatt	aatcctgcat	gctggacatg	tatggtaata	atttctattt	20 40
tgtaccattt	ctttggttaa	ctttagcattg	ttgttgatca	tggatcatac	tccccttggt	21 00
tctttgggtg	agaagggatac	gcagtttgrga	aactccggcg	gctgcgtgctg	gggtttcagt	21 60
cccagctgta	ggcttgtaaa	tacccgcccc	gccaaaccgc	atagagaacg	tggcagcaag	22 20
ctgagggtct	ttgtttgggt	ttattattac	ggtatttttg	tttgtaagtt	aaaaagaaaa	22 80
aaaaaaagaa	aaagttccgg	gcattttgca	tcagaaaaca	actttgtctt	ggggcacact	23 40
tgggaagtgc	atgttttctt	tccttccctt	atccccattc	ggtcctcttt	ttcctctctc	24 00
gcttttagttt	tcaaccttgt	tggtgctgag	agagagaacc	gagagggtccc	agtacaaggg	24 60
cagggcaggg	caggggaagct	gccaaagctcc	gcaccccgag	ggagtgttct	ggactacagc	25 20
cttgtcttat	ggtcaaattg	atacccttaa	taagaaagga	aaggaaagga	aaacagatcc	25 80

```

tcccctctgc tttttattgt aaccagaatc accctgaggt cccttctgaa ccctc tgggc 2640
ctgcgctaata ttaggagcc acagcgctcc tagggtgaga ggcttagcca tccct gaccc 2700
tggcagtgca ctggtaagca gacactgcac tgaaccaact gctatgctca gaatg tacca 2760
gaaacccaaa cattggcaag taattttgca actttcaagt gcgttcttta gaccaatgca 2820
ttgcgtttct tcccctgctt ttgagatagt aggaagagtt cttgggtgggtg tcccc cccct 2880
tcaattcttc agttgtatag tagttatagg gaagatatgg gtgtttttct ttattattac 2940
tttttttttt ctgcaggtca gtaaaaggat ttaagttgca ctgacaaaaa taccaaaata 3000
aaagtgtatt tttaagttcc catttgaaat tgctggcgct gctggccgga tgcat ttttg 3060
agtttgtatt agttgataaa ttaacagtaa taacaagatt gtatgaaccg catggtgctt 3120
gcagttttaa atattgtgga tatttgtcct gcatcagaaa cgagcttttg tttttacaga 3180
ttcaactgtg ttgaaatcaa acctgccgca acagaaattg tttttatttc atgtaaaata 3240
agggatcaat ttcaaaccct gcttatgata tgaaaatatt aaaacctagt ctattgtagt 3300
tttattcaga ctggtttctg ttttttggtt attaaaatgg tttcctattt tgcttattaa 3360
aaaaaaaaaa aaaaaaaa 3378

```

<210> 182

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005461

<400> 182

```

atttgtcctg catcagaaac gagctttggt ttttacagat tcaactgtgt tgaaa tcaaa 60

```

<210> 183

<211> 597

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005532

<400> 183

```

agctgaagtt gaggatctct tactctctaa gccacggaat taacccgagc aggca tggag 60
gcctctgctc tcacctcatc agcagtgacc agtgtggcca aagtggtcag ggtgg cctct 120
ggctctgccg tagttttgcc cctggccagg attgctacag ttgtgattgg aggagtgtgtg 180
gccatggcgg ctgtgcccat ggtgctcagt gccatgggct tcaactgccc gggaa tccc 240
tcgtcctcca tagcagccaa gatgatgtcc gcggcgccca ttgccaatgg ggtgaggatt 300
gcctcgggca gccttgtggg tactctgcag tcaactgggag caactggact ctccgggattg 360
accaagttca tcttgggctc cattgggtct gccattgcgg ctgtcattgc gaggt tctac 420
tagctccctg ccctcgcgcc tgcagagaag agaaccatgc caggggagaa ggcac ccagc 480
catcctgacc cagcgaggag ccaactatcc caaatatacc tgggtgaaat atacc aaatt 540
ctgcatctcc agaggaaaat aagaaataaa gatgaattgt tgcaactctt aaaaaa 597

```

<210> 184

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005532

<400> 184

```

agccaactat cccaaatata cctgggtgaa atataccaaa ttctgcatct ccagaggaaa 60

```

<210> 185

<211> 1661

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005566

<400> 1 85

tgctgcagcc	gctgcccgcg	attccggatc	tcattgccac	gcgcccccca	cgaccgcccg	60
acgtgcattc	ccgattccct	ttgggtccaa	gtccaatatg	gcaactctaa	aggatcagct	120
gatttataat	cttctaaagg	aagaacagac	cccccagaat	aagattacag	ttgttggggg	180
tggtgctgtt	ggcatggcct	gtgccatcag	tatcttaatg	aaggacttgg	cagatgaact	240
tgctcttggt	gatgtcatcg	aagacaaatt	gaagggagag	atgratggatc	tccaacatgg	300
cagccttttc	cttagaacac	caaagattgt	ctctggcaaa	gactataatg	taactgcaaa	360
ctccaagctg	gtcattatca	cggctggggc	acgtcagcaa	gagggagaaa	gccgtcttaa	420
tttggtccag	cgtaacgtga	acatatttaa	attcatcatt	cctaatgttg	taaaatacag	480
cccgaactgc	aagttgctta	ttgtttcaaa	tccagtggat	atcttgacct	acgtggcttg	540
gaagataagt	ggttttccca	aaaaccgtgt	tattggaagt	ggttgcaatc	tggattcagc	600
ccgattccgt	tacctgatgg	gggaaaggct	gggagttcac	ccaattaagct	gtcatgggtg	660
ggtccttggg	gaacatggag	attccagtgt	gcctgtatgg	agtggaaatga	atgttgctgg	720
tgtctctctg	aagactctgc	accagatttt	agggactgat	aaagataagg	aacagtggaa	780
agaggttcac	aagcaggtgg	ttgagagtgc	ttatgaggtg	atcaaaactca	aaggctacac	840
atcctgggct	attggactct	ctgtagcaga	tttggcagag	agtataatga	agaatcttag	900
gcgggtgcac	ccagtttcca	ccatgattaa	gggtctttac	ggaataaagg	atgatgtctt	960
ccttagtggt	ccttgcatct	tgggacagaa	tggaaatctca	gaccttggtga	aggtgactct	1020
gacttctgag	gaagaggccc	gtttgaagaa	gagtgcagat	acactttggg	ggatccaaaa	1080
ggagctgcaa	ttttaaaagtc	ttctgatgtc	atatcatttc	actgtctagg	ctacaacagg	1140
attctaggtg	gaggttgtgc	atgttgtcct	ttttatctga	tctgtgatta	aagcagtaat	1200
attttaagat	ggactgggaa	aaacatcaac	tcctgaagtt	agaaataaga	atggtttgta	1260
aaatccacag	ctatatcctg	atgctggatg	gtattaatct	tgtgtagtct	tcaactgggt	1320
agtgtgaaat	agttctgcca	cctctgacgc	accactgcc	atgctgtacg	tactgcattt	1380
gccccttgag	ccaggtggat	gtttaccgtg	tgttatataa	cttcctggct	ccttcaactga	1440
acatgcctag	tccaacattt	tttcccagtg	agtcacatcc	tgggatccag	tgtataaatc	1500
caatatcatg	tcttgtgcat	aattcttcca	aaggatctta	ttttgtgaac	tatatcagta	1560
gtgtacatta	ccatataatg	taaaaagatc	tacatacaaa	caatgcaacc	aactatccaa	1620
gtgttatacc	aactaaaacc	cccaataaac	cttgaacagt	g	1661	

<210> 186

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005566

<400> 186

catcaactcc	tgaagttaga	aataagaatg	gtttgtaaaa	tccacagcta	tatcctgatg	60
------------	------------	------------	------------	------------	------------	----

<210> 187

<211> 2993

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005689

<400> 187

gggctgcag	ttggcagaag	ggtcccgggc	ccagagccag	cgggggcctg	ctgagacggc	60
gtacgtgcc	tgcgtgagtg	cgtggcgggc	gcgcgtgcgc	tagggggagt	ggcggtgagg	120
cctggtccac	gtgcgtccct	tcccgggacc	cccgcagctt	ggcgcccagc	ggctacgtga	180
gccaaagcac	ccggatgtcc	gcgcccctct	ccgagtgaca	agtcgccggc	tccgggtccc	240
cagtgcccgc	agcctcggcc	ggcgtccacg	cattgccatg	gtgactgtgg	gcaactactg	300
cgaggccgaa	gggcccgtgg	gtccggcctg	gatgcaggat	ggcctgagtc	cctgcttctt	360
cttcacgctc	gtgccctcga	cgcggatggc	tctagggact	ctggccttgg	tgctggctct	420

```

tcctctgcaga cgccggggagc ggcccgcctgg tgctgattcg ctgtctttggg gggccgggccc 480
tcgcatctctt ccctacgtgc tgcagctgct tctggccaca ctccaggcgg cgctgcccct 540
ggccggcctg gctggccggg tgggcactgc ccggggggcc ccactgccaa gctatctact 600
tctggcctcc gtgctggaga gtctggccgg cgctgtggc ctgtggctgc ttgtcgtgga 660
gaggagccag gcacggcagc gtctggcaat gggcatctgg atcaagttca ggcacagccc 720
tggtctcctg ctctcttgga ctgtggcggt tgcagctgag aactttggccc tgggtgtctt 780
gaacagccca cagtgggtgg gggcaagggc agacttgggc caacagggtc agtttagcct 840
gtgggtgctg cgggtatgtg tctctggagg gctgtttgtc ctgggtctct gggcccctgg 900
acttcgtccc cagtcctata cattgcaggt tcatgaagag gaccaagatg tggaaaggag 960
ccagggttcgg tcagcagccc aacagtctac ctggcgagat tttggcagga agctccgct 1020
cctgagtggtc tacctgtggc ctcgaggagg tccagctctg cagctgggtg tgctcatctg 1080
cctggggctc atgggtttgg aacgggcact caatgtgttg gtgcctatat tctataggaa 1140
catctgtgaac ttgctgactg agaaggcacc ttggaactct ctggcctgga ctgttaccag 1200
ttaagctctc ctcaagttcc tccagggggg tggcactggc agtacaggct tcgtgagcaa 1260
cctgcgcacc ttctcttgga tccgggtgca gcagttcacg tctcggcggg tggagctgct 1320
catcttctcc cacctgcacg agctctcact gcgtggcac ctggggcgcc gcacagggga 1380
ggtgctgcgg atcgcggatc ggggcacatc cagtgtcaca ggctggtca gctacctggt 1440
gttcaatgct atccccacg tggccgacat catcattggc atcatctact tcagcatggt 1500
cttcaacgcc tgggtttggc tcattgtgtt cctgtgcag agtctttacc tcacctgac 1560
catctgtggtc actgagtgga gaaccaagtt tcgtcgtgct atgaacacac aggagaacgc 1620
taaccgggca cgagcagtg actctctgct aaacttcgag acggtgaagt attacaacgc 1680
cgagagttac gaagtggaa gctatcgaga ggccatcatc aaatatcagg gtttggagtg 1740
gaagtcgagc gcttcaactg ttttactaaa tcagacccag aacctggtga ttgggctcgg 1800
gctcctcgcc ggctccctgc tttgcgcata ctttgtcact gaggcagaagc tacagggttg 1860
ggactatgtg ctcttttgga cctacattat ccagctgtac atgcccctca attgggtttg 1920
cacctactac aggatgatcc agaccaactt cattgacatg gagaacatgt ttgactgtct 1980
gaaagaggag acagaagtga aggaccttc tggagcaggg ccctctcgct ttcagaaggg 2040
cgttattgag tttgagaacg tgcacttcag ctatgccgat gggcgggaga ctctgcagga 2100
cgtgtctttc actgtgatgc ctggacagac acttgccctg gtgggcccac ctggggcagg 2160
gaagagcaca attttgcgcc tgctgtttcg cttctacgac atcagctctg gctgcatccg 2220
aatagatggg caggacattt cacagggtgac ccaggcctct ctccgggtct acattggagt 2280
tgtgccccaa gacactgtcc tctttaatga caccatcgcc gaccaatatc gttacggccg 2340
tgtcacagct ggggaatgat aggtggaggc tgctgctcag gctgcaggca tccatgatgc 2400
catctatggc ttccctgaag ggtacaggac acagggtggc gaggcggggac tgaagctgag 2460
cggcggggag aagcagcgcg tcgccattgc ccgcaccatc ctcaaggctc cgggcatcat 2520
tctgctggat gaggcaacgt cagcgtgga tacatctaag gagagggcca tccaggcttc 2580
tctggccaaa gtctgtgcca accgcaccac catcgtagtg gcacacaggc tctcaactgt 2640
ggtcaatgct gaccagatcc tcgtcatcaa ggatggctgc atcgtggaga ggggacgaca 2700
cgaggctctg ttgtcccgag gtgggtgtg tgctgacatg tggcagctgc agcagggaca 2760
ggaagaaacc tctgaagaca ctaagcctca gaccatggaa cggtgacaaa agtttggcca 2820
cttccctctc aaagactaac ccagaaggga ataagatgtg tctcctttcc ctggcttatt 2880
tcatcctggt cttggggtat ggtgctagct atggtaaggg aaagggacct ttccgaaaaa 2940
catcttttgg ggaaataaaa atgtggactg tgaaaaaaa aaaaaaaa aaa 2993

```

<210> 188

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005689

<400> 188

ggaaagggac ctttccgaaa aacatctttt ggggaaataa aaatgtggac tgtgaaaaaa 60

<210> 189

<211> 1830

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 189

```

ggggagttga aacctaattt tgtggcgtag cagctatgca gcttgaaatc caagtagcac 60
taaa ttttat tatttcgtat ttgtacaata agcttcccag gagacgtgtc aacatttttg 120
gtga agaact tgaaagactt cttaagaaga aatatgaagg gcaactggtat cctgaaaagc 180
cata caaagg atcggggttt agatgtatac acatagggga gaaagtggac ccagtgattg 240
aaca agcatc caaagagagt ggtttggaca ttgatgatgt tctgtggcaat ctgccacagg 300
atct tagtgt ttggatcgac ccatttgagg ttctttacca aatttggtgaa aagggaccag 360
tgaa ggtgct ttacgtggat gataataatg aaaatggatg tgagttggat aaggagatca 420
aaaa cagctt taaccagag gccaggttt ttatgccc atagtgacca gcctcatcag 480
tgct cagctc tccatcgct ccttttggtc actctgctgc tctaagccct accttcacgc 540
cccggtccac tcagccttta acctttacca ctgccacttt tctgtgccac aagttcggct 600
ctac caaaat gaagaatagt ggccgtagca acaagggtgc acgtacttct cccatcaacc 660
tcgg cttgaa tgtgaatgac ctcttgaagc agaaagccat ctcttcctca atgcactctc 720
tgta tgggct tggccttgggt agccagcagc agccacagca acagcagcag ccagcccagc 780

```

```

cgcc accgcc accaccacca ccacagcagc aacaacagca gaaaacctct gctctttctc 840
ctaa tgccaa ggaattttatt ttctctaata tgcagggtca aggtagtagt accaatggaa 900
tggt cccagg tgacagcccc cttaacctca gtctctcca gtacagtaat gcctttgatg 960
tggt tgcagc ctatggaggc ctcaatgaga agtcttttgt agatggcttg aatttttagct 1020
taaa taacat gcagttattct aaccagcaat tccagcctgt tatggctaac taaaaaaaaaag 1080
aaaa tgtatc gtacaagtta aaatgcacgg gcccaagggg gatttttttt ttacacctct 1140
tgagaatttt tttttttaag cttatagtaa ggatacatte aagcttggtt aaaaaataaa 1200
taat aaaaca tgcattcatt ttcatcttgc aaccaagcac aaagtatttt tatactgact 1260
gtat atttta agtatactc tcagatatgg cctcttacag tatttaagat atagcaagga 1320
catggctgat ttttttttat aaaaattggc actaataagt ggggtttattg gtcttttcta 1380
attgtataat ttaatttagt acaaagtttg taaaatatca gaggatatat atatatgtt 1440
tcta cgacat ggtattgcat ttatatcttt ttactacagt gatctgtgac agcagcagct 1500
tcat gttgta ttttttttac tgaaattgta aaatatccat cttaaagaca tcaactattc 1560
taaaaattgt gtacaggata ttcttttagt ggtggaaatta aaatgtacga atacttgctt 1620
tttcaaaaaa atgtattttc tgttaaaagt ttaaagattt ttgctatata ttatggaaga 1680
aaaa tgtaat cgtaaattt aattttgtac ctatattgtg caatacttga aaaaaacggt 1740
ataa aagtat tttgagtcag tgtcttacat gttaaagagg actgaaatag tttatatata 1800
gttt gtatta aaattcttta aaattaaaaa 1830

```

<210> 190

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 190

```

aaacctctgc tctttctcct aatgccaaagg aatttatattt tcctaatatg caggggtcaag 60

```

<210> 191

<211> 1534

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005804

<400> 191

```

ggaa ggcgcag caactcgtgt ctgagcgccc ggcggaacac cgaagttgga agtgtctctt 60
agca ggcgcgc ggagaagaac ggggagccag catcatggca gaacaggatg tggaaaacga 120
tctt ttggat tacgatgaag aggaagagcc ccaggctcct caagagagca caccagctcc 180

```

```

ccctaagaaa gacatcaagg gatcctacgt ttccatccac agctcttggt tccgggactt 240
tctgctgaag ccggagctcc tgcggggccat cgtggactgt ggctttgagc atccttctga 300
ggtccagcat gagtgcattc cccaggccat cctgggcatg gacgtcctgt gccaggccaa 360
gtccgggatg ggcaagacag cggctcttcgt gctggccacc ctacagcaga ttgagcctgt 420
caacggacag gtgacgggtcc tggatcatgt ccacacgagg gagctggcct tccagatcag 480
caaggaatat gagcgctttt ccaagtacat gcccagcgtc aagggtgtctg tgttcttcgg 540
tggctctctc atcaagaagg atgaagaagt gttgaagaag aactgtcccc atgtcgtggt 600
ggggaccccc ggccgcatcc tggcgctcgt gcggaatagg agcttcagcc taaagaatgt 660
gaagcacttt gtgctggacg agtgtgacaa gatgctggag cagctggaca tgcggcggga 720
tgtgcaggag atcttccgcc tgacaccaca cgagaagcag tgcattgatgt tcagcgccac 780
cctgagcaag gacatccggc ctgtgtgcag gaagttcatg caggatccca tggagggtgtt 840
tgtggacgac gagaccaagc tcacgctgca cggcctgcag cagtactacg tcaaactcaa 900
agacagttag aagaaccgca agctctttga tctcttggtg gtgctggagt ttaaccagg 960
gataatcttc gtcaagtcag tgcagcgctg catggccctg gccagctcc tcgtggagca 1020
gaacttcccc gccatcgcca tccaccgggg catggcccag gagggagcgcc tgtaacgcta 1080
tcagcagttc aaggatttcc agcggcggat cctggtggcc accaatctgt ttggccgggg 1140
gatggacatc gagcgagtca acatcgtctt taactacgac atgcctgagg actcggacac 1200
ctacctgtac cgggtggccc gggcgggtcg ctttggcacc aaaggcctag ccatcacttt 1260
tgtgtctgac gagaatgatg ccaaaatcct caatgacgtc caggaccggt ttgaagttaa 1320
tgtggcagaa cttccagagg aaatcgacat ctccacatac atcgagcaga gccggtaacc 1380
accacgtgcc agagccgccc acccgagacc gcccgcatgc agcttcacct cccctttcca 1440
ggcgccactg ttgagaagct agagattgta tgagaataaa cttgttatta tggaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1534

```

<210> 192
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005804

```

<400> 192
gttgagaagc tagagattgt atgagaataa agtgttatta tgaaatgaag aagcctcacc 60

```

<210> 193
 <211> 1416

<212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 1416
 <223> n = a, c, g, or t

<300>
 <308> NM_005945

```

<400> 193
aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttggtg 60
cgaacgccat tctgggggtg tcccttgccg tctgcaaagc tgggtgccgtt gagaaggggg 120
tccccgttac cgccacatcg cgtacttggc tggcaacttc gaagtcattc tgccagtccc 180
ggcgttcaag tgtcatcatc aatggcgggt ctcatgctgg caacaagctg gccatgcaga 240
gtctgtcctc ccagtcgggtg cagcaaactc agggaaagcca tgcgcgattg gagcagagg 300
ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgcca ccaatgtggg 360
gatttgcgcg ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
actgctattg gaaagcctgg cctacactgt aaagggtggtc atggcatgga cgtagcggcc 480
tccgagttct tcaggtcagg gaactatgac ctggacttca agtctcccga tgacccagc 540
aggtacatct cgctgacca gctggctgac ctgtacaagt ccttcatcaa ggactacca 600
gtggtgtcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660

```



```

cagtgcagga atccaggtag tgggggggatg actcacagtg accaaccctaa agaggatcgc 720
caaggcgctga acgagaagtc ctgcaactgc ctcttctca aagtcaacca gattggctcc 780
gtgaccgagt ctcttcaggc gtgcaagctg gccaggcca atgggttggg cgatcatggtg 840
tctcatcggt cgggggagac tgaagatacc ttcacgctg acctgggtgt ggggctgtgc 900
actggggcag atcaagactg gtgccccttg ccgatcacgc gcttggccaa gtacaaccag 960
ctcctcagaa ttgaagagga gctgggcagc aaggctaagt ttgcccggcag gaacttcaga 1020
aaccctcttg ccaagtaagc tgtgggcagg caagccttcg gtcacctgtt ggctacagac 1080
ccctccctg gtgtcagctc aggcagctcg agggcccgga ccaacacttg caggggtccc 1140
tgctagttag cgcccaccgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagctccctg gaagccctgt tggcagctct agctttgcag ttgtgttaatt ggcccaagtc 1260
attgtttttc tcgccttact ttccaccaag tgtctagagt catgtgagcc tngtgtcatc 1320
tccgggggtg ccacaggcta gatccccggt gggtttgtgc tcaaaataaa aagcctcagt 1380
gaccatgaa aaaaaaaaaa gaattccgga attccg 1416

```

<210> 194

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005945

<400> 194

```

ttgtgttaatt ggcccaagtc attgtttttc tcgccttact ttccaccaag tgtctagagt 60

```

<210> 195

<211> 961

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006014

<400> 195

```

ggcgaccacg gtgtcttcaa aagccccgct aggggttggt tcttggggcc ggaccgactg 60
tgggtcagtt tgcaccagcg ctctggaatc gagttacgcg cgaaggggca gagtttctgg 120
aggaaaccgc agcctctcaa ccgctgaccg ggtctcagaa ggcccccggc agggcgcgtt 180
ggcggggaact gaccacgcgc cagtcaggct ctccagggac ctgctgcaggc gcgtgtgggc 240
ggagtcgtgc gcagggggcg gggcttcggg aaggagccac agagagggcg gggcgtagga 300
cctgcgcttc gggggtggag tcggagcggc gggcgggcg tcatgaggga cgcggatgca 360
gacgcaggcg gaggcgctga cggcggggat ggccgggggt gccacagctg ccgcgggggc 420
gtggacacag ccgcagctcc ggccgggtga gctccccag cgcacgcgcc aggtccgggc 480
agagacgccg cgtctgcggc cagggggtca cgaatgcggc cgcacatatt caccctcagc 540
gtgcctttcc cgaccctt ggaggcgga atcgccatg ggtccctggc accagatgcc 600
gagccccacc aaagggtggt tgggaaggat ctacagtgga tgggcaggat cctggctcgtc 660
cgctggaaag ctgaagactg tcgcctgctc cgaatttccg tcatcaactt tcttgaccag 720
ctttccctgg tggtagcgag catgcagcgc tttgggcccc ccgttttccg ctaagcctgg 780
cctgggcaaa tggagcgagg tcccactttg cgtctccttg taggcagtgc gtccatcctt 840
ccctagggca ggaattccca cagttgctac tttcctggga gggcctcatg ttttatctgg 900
ttcttaaatg tttgttacta cagaaaataa aactgcgcta ctaaaaaaaaa aaaaaaaaaa 960
a 961

```

<210> 196

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006014

<400> 196

ggcctcatgt tttatctggt tcttaaagt ttgttactac agaaaataaa actgaggtat 60

<210> 197
 <211> 1648
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 197
 atgcgggaga tctgtcacat ccaggccggc cagtgcggca accagatcgg ggccaagtgc 60
 tgggaagtca tcagtgatga gcatggcatc gacccagcgg gcaactacgt gggcgactcg 120
 gacttgcagc tggagcggat cagcgtctac tacaacgagg cctcttctca caagtacgtg 180
 cctcgagcca ttctggtgga cctggaaccc ggaaccatgg acagtgtccg ctgaggggcc 240
 tttggacatc tcttcaggcc tgacaatttc atctttgggc agagtggggc cggcaacaac 300
 tgggccaagg gtcactacac ggagggggcg gagctggtgg attcggtcct ggatgtggtg 360
 cggaaaggagt gtgaaaactg cgactgcctg cagggcttcc agc tgacca ctcgctgggg 420
 ggggggacgg gctccggcat gggcacgttg ctcatcagca aggtgcgtga ggagtatccc 480
 gaccgcatca tgaacacctt cagcgtcgtg ccctcaccca aggtgtcaga cacggtggtg 540
 gaaccctaca acgccacgct gtccatccac cagctggtgg aaaacacgga tgaaacctac 600
 tgcacgaca acgaggcgct ctacgacatc tgcttccgca ccc tcaagct ggccacgccc 660
 acctacgggg acctcaacca cctggtatcg gccaccatga gcggagtcac cacctccttg 720
 cgcttccggg gccagctcaa cgctgacctg cgcaagctgg ccgtcaacat ggtgcccttc 780
 ccggcctgc acttcttcat gccgggcttc gccccctca cca ggcgggg cagccagcag 840
 taccgggccc tgaccgtgcc cgagctcacc cagcagatgt tcgatgccaa gaacatgatg 900
 gcgcctgcg acccgcgcca cggccgctac ctgacgggtg cca ccgtgtt ccggggccgc 960
 atgtccatga aggaggtgga cgagcagatg ctggccatcc aga gcaagaa cagcagctac 1020
 ttcgtggagt ggatcccca caacgtgaag gtggcgtgt gtgacatccc gccccgcggc 1080
 ctcaagatgt cctccacctt catcggaac agcacggcca tccaggagct gttcaagcgc 1140
 atctccgagc agttcacggc catgttccgg cgcaaggcct tcc tgcactg gtacacgggc 1200
 gagggcatgg acgagatgga gttcaccgag gccgagagca aca tgaacga cctggtgtcc 1260
 gagta ccagc agtaccagga cgccacggcc gaggaagagg gcgagatgta cgaagacgac 1320
 gaggaaggagt cggaggccca gggccccaag tgaaactgct cgc agctgga gtgagaggca 1380
 ggtggcggcc ggggccgaag ccagcagtgt ctaaaacccc gga gccatct tgctgccgac 1440
 accctgcttt ccccatcgcc ctagggtctc cttgccgccc tcc tgcagta tttatggcct 1500
 cgctctcccc cacctaggcc acgtgtgagc tgctcctgtc tctgtcttat tgcagctcca 1560
 ggcctgacgt tttacggttt tggttttttac tggtttgtgt tta tattttc ggggatactt 1620
 aataaatcta ttgctgtcag ataccctt 1648

<210> 198
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 198
 ttttactgg tttgtgttta tttttcggg gatacttaat aaatctattg ctgtcagata 60

<210> 199
 <211> 3074
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006096

<400> 199
 aacaaacctc gcctggctcc cagctggtgc tgaagctcgt cagttcacca tccgcctcg 60

gcttccgcg	ggcgctgggc	cgccagcctc	ggcacogtcc	tttccctttct	ccctcgcggt	120
aggcaggtga	cagcagggac	atgtctcggg	agatgcagga	tgtagacctc	gctgaggtga	180
agcctttggt	ggagaaaggg	gagaccatca	ccggcctcct	gcaagagttt	gatgtccagg	240
agcaggacat	cgagacttta	catggctctg	ttcacgtcac	gctgtgtggg	actcccaagg	300
gaaaccggcc	tgtcatcctc	acctaccatg	acatcggcat	gaaccacaaa	acctgctaca	360
acccccctct	caactacgag	gacatgcagg	agatcaccca	gcacttttgc	gtctgccacg	420
tggacgcccc	tggccagcag	gacggcgag	cctccttccc	cgcagggtag	atgtacccct	480
ccatggatca	gctggctgaa	atgcttctct	gagtccctca	acagttttgg	ctgaaaagca	540
ttattggcat	gggaacagga	gcaggcgctc	acatcctaac	tcgattttgt	ctaaacaacc	600
ctgagatggt	ggagggcctt	gtccttatca	acgtgaaccc	ttgtgcggaa	ggctggatgg	660
actgggcccc	ctccaagatc	tcaggatgga	cccaagctct	gccggacatg	gtggtgtccc	720
accttttttg	gaaggaagaa	atgcagagta	acgtggaagt	ggtccacacc	taccgccagc	780
acattgtgaa	tgacatgaac	cccggcaacc	tgcacctgtt	catcaatgcc	tacaacagcc	840
ggcgcgacct	ggagattgag	cgaccaatgc	cggaacacca	cacagtcacc	ctgcagtgcc	900
ctgctctgtt	ggtggttggg	gacagctcgc	ctgcagtgga	tgccgtggtg	gagtgcact	960
caaaatttga	cccaacaaa	accactctcc	tcaagatggc	ggactgtggc	ggcctcccgc	1020
agatctccca	gcccggcaag	ctcgtgagg	ccttcaagta	cttcgtgcag	ggcatgggat	1080
acatgccctc	ggctagcatg	acccgcctga	tgcggtcccg	cacagcctct	ggttccagcg	1140
tcactttctc	ggatggcacc	cgcagccgct	cccacaccag	cagagggcacc	cgaagccgct	1200
cccacaccag	cgagggcacc	cgcagccgct	cgcacaccag	cagagggggc	cacctggaca	1260
tcacccccaa	ctcgggtgct	gctgggaaca	gcccggggcc	caagtccatg	gaggtctcct	1320
gctagggcgg	ctgccagctc	gcccggcccg	gactctgata	ctctgtagtg	ccccctcctc	1380
cccggccccc	tttcgcccc	tgccctgccat	actgcgccta	actcgggtatt	aatccaaagc	1440
ttatttttga	agagtgaagt	ctggtggaga	caaagtgggt	ctattacgtg	ggtgccctct	1500
ccaaaggcgg	ggtggcggtg	gaccaaagga	aggaagcaag	catctccgca	tcgcatcctc	1560
ttccattaac	cagtggccgg	ttgccactct	cctcccctcc	ctcagagaca	ccaaactgcc	1620
aaaaacaaga	cgcgtagcag	cacacacttc	acaaagccaa	gcctaggccg	ccctgagcat	1680
cctggttcaa	acgggtgcct	ggtcagaagg	ccagccgccc	acttcccgtt	tcctctttaa	1740
ctgaggagaa	gctgatccag	tttccggaaa	caaaatcctt	ttctcatttg	gggagggggg	1800
taatagtga	atgcaggcac	ctctttttaa	caggcaaaac	aggaaggggg	aaaagggtgg	1860
attcatgtcg	aggctagagg	catttggaac	aacaaatcta	cgtagttaac	ttgaagaaac	1920
cgatttttaa	agttggtgca	tctagaaaag	tttgaatgca	gaagcaaaac	agcttgattt	1980
ttctagcatc	ctcttaaatg	gcagcaaaa	caggcgacaa	aatctcctgg	ctttacagac	2040
aaaaatatct	cagcaaacgt	tgggcatcat	ggtttttgaa	ggcttttagt	ctgctttctg	2100
cctctcctcc	acagccccaa	cctcccaccc	ctgatacatg	agccagtgat	tattcttggt	2160
cagggagaag	atcatttaga	tttgttttgc	attccttaga	atggagggca	acattccaca	2220
gctgccctgg	ctgtgatgag	tgtccttgca	ggggccggag	taggagcact	ggggtggggg	2280
tgggaattgg	gttactcgat	gtaagggatt	ccttggtggt	gtgttgagat	ccagtgcagt	2340
tgtgatttct	gtggatccca	gcttggttcc	aggaattttg	tgtgattggc	ttaaatccag	2400
ttttcaatct	tcgacagctg	ggctggaaag	tgaactcagt	agctgaacct	gtctgacccg	2460
gtcacgttct	tggatcctca	gaactccttg	ctcttgtcgg	ggtgggggtg	ggaactcacg	2520
tggggagcgg	tggctgagaa	aatgtaagga	ttctggaata	catattccat	gggactttcc	2580
ttccctctcc	tgcttctctc	tttctgtctc	cctaaccttt	cggcgaatgg	ggcagcacca	2640
ctgacgtttc	tgggcggcca	gtgcggctgc	caggttcctg	tactactgcc	ttgtactttt	2700
catttttggt	caccgtggat	tttctcatag	gaagtttggt	cagagtgaat	tgaatatatt	2760
aagtcagcca	ctgggacccg	aggatttctg	ggaccccgca	gttggggagga	ggaagtagtc	2820
cagccttcca	ggtggcgtga	gaggcaatga	ctcgttacct	gccgcccata	accttggagg	2880
ccttccctgg	ccttgagtag	aaaagtcggg	gacgggggca	agagaggctg	agtacggatg	2940
ggaaactatt	gtgcacaagt	ctttccagag	gagtttctta	atgagatatt	tgtattttatt	3000
tccagaccaa	taaatttgta	actttgcagc	ggaaaaaaa	aaaaaaaaa	aaaaaaaaa	3060
aaaaaaaaa	aaaa	3074				

<210> 200

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006096

<400> 200
 gagtaCggat gggaaactat tgtgcacaag tctttccaga ggagtttctt aatgagatat 60

<210> 201
 <211> 2148
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006115

<400> 201
 gcttcaggggt acagctcccc cgcagccaga agccgggcct gcagcgcctc agcaccgctc 60
 cgggacacccc caccgcgttc ccaggcgtga cctgtcaaca gcaacttcgc ggtgtggtga 120
 actctctgag gaaaaacat tttgattatt actctcagac gtgcgtggca acaagtgact 180
 gagacctaga aatccaagcg ttggagggtcc tgaggccagc ctaagtcgct tcaaaatgga 240
 acgaaggcgt ttgtgggggt ccattcagag ccgatacatc agcatgagtg tgtggacaag 300
 cccacggaga cttgtggagc tggcagggca gagcctgctg aaggatgagg ccctggccat 360
 tgccgccctg gagttgctgc ccaggagagc cttcccgcca ctcttcatgg cagcctttga 420
 cgggagacac agccagaccc tgaaggcaat ggtgcaggcc tggcccttca cctgcctccc 480
 tctgggagtg ctgatgaagg gacaacatct tcacctggag accttcaaag ctgtgcttga 540
 tggacttgat gtgctccttg cccaggaggt tcgccccagg aggtggaaac ttcaagtgtc 600
 ggatttacgg aagaactctc atcaggactt ctggactgta tggctctggaa acagggccag 660
 tctgtactca tttccagagc cagaagcagc tcagcccatg acaaagaagc gaaaagtaga 720
 tggtttgagc acagaggcag agcagccctt cattccagta gaggtgctcg tagacctgtt 780
 cctcaaggaa ggtgcctgtg atgaattgtt ctctacctc attgagaaaag tgaagcgaaa 840
 gaaaaaatgta ctacgcctgt gctgtaagaa gctgaagatt tttgcaatgc ccatgcagga 900
 tatcaagatg atcctgaaaa tgggtgcagct ggactctatt gaagatttgg aagtgaactg 960
 tacctggaag ctaccacac tggcgaaatt ttctccttac ctgggcccaga tgattaatct 1020
 gcgtagactc ctctctccc acatccatgc atcttcctac atttccccgg agaaggaaga 1080
 gcagtatatc gccagttca cctctcagtt cctcagtcctg cagtgcctgc aggtctctta 1140
 tgtggactct ttatTTTTTcc ttagaggccg cctggatcag ttgctcaggc acgtgatgaa 1200
 cccttgga accctctcaa taactaactg ccggctttcg gaaggggatg tgatgcatct 1260
 gtcccagagt cccagcgtca gtcagctaag tgtcctgagt ctaagtgggg tcatgctgac 1320
 cgatgtaagt cccagcctcc tccaagctct gctggagaga gcctctgcca ccctccagga 1380
 cctggctctt gatgagtggt ggatcacgga tgatcagctc cttgccctcc tgccttccct 1440
 gagccactgc tcccagctta caaccttaag cttctacggg aatttccatct ccatactctgc 1500
 cttgcagagt ctctgcagc acctcatcgg gctgagcaat ctgacccacg tgctgtatcc 1560
 tgtccccctg gagagttatg aggacatcca tggtaacctc caacctggaga ggcttgccct 1620
 tctgcatgcc aggtcagggg agttgctgtg tgagttgggg cggcccagca tgggtctggct 1680
 tagtgccaac ccctgtcctc actgtgggga cagaaccttc tatgaccggg agcccatcct 1740
 gtgccctgt ttcatgccta actagctggg tgcacatatc aaatgcttca ttctgcatac 1800
 ttggacata aagccaggat gtgcatgcat cttgaagcaa caaagcagcc acagtttcag 1860
 acaaa tgttc agtgtagtg aggaaaacat gttcagtgag gaaaaaacat tcagacaaat 1920
 gttcagtga gaaaaaagg ggaagttggg gataggcaga tgttgacttg aggagttaat 1980
 gtgatctttg gggagataca tcttatagag ttagaaatag aa tctgaatt tctaaaggga 2040
 gattc tggct tgggaagtac atgtaggagt taatccctgt gtgactgtt gtaaagaaac 2100
 tgttgaaaat aaagagaagc aatgtgaagc aaaaaaaaaa aaaaaaaa 2148

<210> 202
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006115

<400> 202
 tggggagata catcttatag agttagaaat agaacttgaa tt tctaaagg gagattcttg 60

<210> 203
 <211> 1051
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006332

<400> 203
 gga ccg ccgc ctggttaaag gcgcttattt cccagggcagc cgctgcagtc gccacacctt 60
 tgc ccctgct gcgatgacct tgcgccact tctgctgttc ctgccaccgc tgcctgctgct 120
 gctggacgtc cccacggcgg cggtgcaggc gtccctctctg caagcggttag acttcttttg 180
 gaa tgggcca ccagttaact acaagacagg caatctatac ctgcggggggc ccctgaagaa 240
 gtc caatgca ccgcttgta atgtgacct ctactatgaa gcactgtgcg gtggctgccg 300
 agc cttcctg atccgggagc tcttcccaac atggctgttg gtcattggaga tcttcaatgt 360
 cac gctggtg ccctacggaa acgcacagga acaaaatgtc agtggcagggt gggagttcaa 420
 gtg ccagcat ggagaagagg agtgcaaatt caacaagggtg gaggcctgcg tgttggtatga 480
 act tgacatg gagctagcct tcttgacct tgtctgcatt gaagagtttg aggacatgga 540
 gaga agtctg ccactatgcc tgcagctcta cggcccagggt ctgtcgccag acactatcat 600
 ggagtgtgca atgggggacc gcggcatgca gctcatgcac gccaacgccc agcggacaga 660
 tgc tctccag ccaccacacg agtatgtgcc ctgggtcacc gtcaatggga aacccttgga 720
 aga tcagacc cagctcctta cccttgctctg ccagttgtac caggggcaaga agccggatgt 780
 ctg cccttcc tcaaccagct ccctcaggag tgtttgcttc aagtgtatggc cgggtgagctg 840
 cgg agagctc atggaaggcg agtgggaacc cggctgcctg cctttttttc tgatccagac 900
 cct cggcacc tgctacttac caactggaaa attttatgca tcccatgaag ccagataca 960
 caaaattcca ccccatgatc aagaatcctg ctccactaag aatggtgcta aagtaaaact 1020
 agt ttaataa gcaaaaaaaaa aaaaaaaaaa a 1051

<210> 204
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006332

<400> 204
 aaattccacc cctagatcaa gaatcctgct ccactaagaa tgggtgctaaa gtaaaactag 60

<210> 205
 <211> 1714
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006417

<400> 205
 ggggcatttt gtgcctgcct agctatccag acagagcagc taccctcagc tctagctgat 60
 actacagaca gtacaacaga tcaagaagta tggcagtgac aactcgtttg acacggttgc 120
 acgaaaagat cctgcaaaat cattttggag ggaagcggct tagccttctc tataagggta 180
 gtgtccatgg attccgtaat ggagttttgc ttgacagatg ttgtaataca gggcctactc 240
 taa cagtgat ttatagttaa gatcatatta ttggagcata tgcggaagag agttaccagg 300
 aaggaaagta tgcttccatc atcctttttg cacttcaaga tactaaaatt tcagaatgga 360
 aac taggact atgtacacca gaaacactgt tttgttgatg tggtacaaaa tataactccc 420
 caa ctaattt ccagatagat ggaagaaata gaaaagtgat tatggactta aagacaatgg 480
 aaaatcttgg acttgctcaa aattgtacta tctctattca ggattatgaa gtttttcgat 540
 gcgaagattc actggatgaa agaaagataa aaggggtcat tgagctcagg aagagcttac 600
 tgt ctgcctt gagaacttat gaaccatatg gatccctggg tcaacaaata cgaattctgc 660
 tgc tgggtcc aattggagct ggaagtcga gctttttcaa ctcaagtgagg tctgttttcc 720

```

aagggcatgt aacgcatacag gcttttggtgg gcactaatac aactgggata tctgagaagt 780
ataggacata ctctattaga gacgggaaag atggcaaata cctgccgttt attctgtgtg 840
actcactggg gctgagtgag aaagaaggcg gcctgtgcag ggatgacata ttctatatct 900
tgaacggtaa cattcgtgat agataccagt ttaatcccat ggaatcaatc aaattaaatc 960
atcatgacta cattgattcc ccatcgctga aggacagaaat tcattgtgtg gcatttgtat 1020
ttgatgccag ctctattcaa tactttctct ctcagatgat agtaaagatc aaaagaattc 1080
gaagggagtt ggtaaacgct gaccttatag aaatagagag atgtgagcct gtgaggtcca 1140
tggatttgat tacaaaagggt gaccttatag aaatagagag atgtgagcct gtgaggtcca 1200
agctagagga agtccaaaga aaacttggat ttgctcttct tgacatctcg gtggttagca 1260
attattcctc tgagtgggag ctggaccctg taaaggatgt tctaattctt tctgctctga 1320
gacgaatgct atgggctgca gatgacttct tagaggattt gccttttgag caaataggga 1380
atctaaggga ggaaattatc aactgtgcac aaggaaaaaa atagatatgt gaaaggttca 1440
cgtaaatctc ctcacatcac agaagattaa aattcagaaa ggagaaaaca cagaccaaag 1500
agaagtatct aagaccaaag ggatgtgttt tattaatgtc taggatgaag aaatgcatag 1560
aacattgtag tacttgtaaa taactagaaa taacatgat tagtcataat tgtgaaaaat 1620
agtaataatt tttcttggat ttatgttctg tatctgtgaa aaaataaatt tcttataaaa 1680
ctcggaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1714

```

<210> 206

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006417

<400> 206

```

atgacatatt ctatatcttg aacggtaaca ttcgtgatag ataccagttt aatcccatgg 60

```

<210> 207

<211> 3791

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006461

<400> 207

```

acagacggcg ggtgaacatg gcgtcctcga cttggtctga gacgtgatag gcctgccttc 60
tggttgaaga tgtggcgagt gaaaaaactg agcctcagc c tgcgccttc gccccagacg 120
ggaaaaccat ctatgagaac tcctctccgt gaacttacc c tgcagcccg tgccctcacc 180
acctctggaa aaagatcccc cgcttgctcc tcgctgacc c catcactgtg caagctgggg 240
ctgcaggaag gcagcaacaa ctcgctcca gtggattttg taaataacaa gaggacagac 300
ttatcttcag aacatttcag tcattcctca aagtggttag aaacttgtca gcatgaatca 360
gatgagcagc ctctagatcc aattcccca attagctcta ctctaaaac gtctgaggaa 420
gcagtagacc cactgggcaa ttatatgggt aaaaccatcg tccttgtacc atctccactg 480
gggcagcaac aagacatgat atttgaggcc cgtttagata ccatggcaga gacaaacagc 540
atatctttta atggaccctt gagaacagac gatctggtga gagaggaggt ggcaccctgc 600
atgggagaca ggttttcaga agttgctgct gtatctgaga aacctatctt tcaggaatct 660
ccgtcccatc tcttagagga gtctccacca aatccctgt t ctgaacaact acattgctcc 720
aaggaaagcc tgagcagtag aactgaggct gtgcgtgagg acttagtacc ttctgaaagt 780
aacgccttct tgccttctc tgttctctgg ctttccctt caactgcctt ggcagcagat 840
ttccgtgtca atcatgtgga ccagaggag gaaattgtag agcatggagc tatggaggaa 900
agagaaatga ggtttcccac acatcctaag gagtctgaaa cagaagatca agcacttgct 960
tcaagtgtgg aagatattct gtccacatgc ctgacacca atctagtaga aatggaatcc 1020
caagaagctc caggcccagc agtagaagat gttggttaga ttcttggctc tgatacagag 1080
tcttggatgt cccactggc ctggctggaa aaagggtgaa atacctccgt catgctggaa 1140
aatctccgcc aaagcttatc ccttccctcg atgcttcggg atgctgcaat tggcactacc 1200
cctttctcta ctgctcggt ggggacttgg tttactcct cagcaccaca ggaaaagagt 1260
acaaacacat ccagacagg cctggttggc accaagcaca gtacttctga gacagagcag 1320
ctcctgtgtg gccggcctcc agatctgact gccttgtctc gacatgactt ggaagataac 1380

```

```

ctgctgagct ctcttgtcat tgtggagttt ctctcccgcc agcttcggga ctggaagagc 1440
cagctggctg tccctcacc cagaaaccag gacagtagca cacagactga cacatctcac 1500
agtgggataa ctaataaaact tcagcatctt aaggagagcc atgagatggg acaggcccta 1560
cagcaggcca gaaatgtcat gcaatcatgg gtgcttatct ct aaagagct gatatccttg 1620
cttcacctat ccctgttgca tttagaagaa gataagacta ctgtgaatca ggagtctcgg 1680
cgtgcagaaa cattggctctg ttgctgtttt gatttgctga agaaattgag ggcaaagctc 1740
cagagcctca aagcagaaaag gtaggaggca aggcacagag aggaaatggc tctcagaggc 1800
aagratgagg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcatcagc 1860
cagctggaac aggacctagc atccatgagg gaattcagag gccttctgaa ggatgccag 1920
acccaactgg tagggcttca tgccaagcaa gaagagctgg tt cagcagac agtgagtctt 1980
acttctacct tgcaacaaga ctggagggtcc atgcaactgg at tatacaac atggacagct 2040
ttgctgagtc ggtcccgcac actcacagag aaactcacag tc aagagcca gcaagccctg 2100
caggaacgtg atgtggcaat tgaggaaaag caggaggttt ct aggggtgct ggaacaagtc 2160
tctgcccagt tagaggagtg caaaggccaa acagaacaac tgagagtggga aaacattcgt 2220
ctagcaacag atctccgggc tcagttgcag attctggcca acatggacag ccagctaaaa 2280
gagctacaga gtcagcatag ccattgtgccc caggacctgg ct atgaagga tgagttactc 2340
tgccagctta ccagagcaa gctggtcaat tgaggagcag gctgctcaat gc gtaaagga agagatggca 2400
ctaaaacaca tgcaggcaga actgcagcag caacaagctg tc ctggccaa agaggtgagg 2460
gacctgaaag agaccttgga gtttgcagac caggagaatc aggttgctca cctggagctg 2520
ggctcaggttg agtgtcaatt gaaaaccaca ctggaagtgc tc cgggagcg cagcttgagc 2580
tgtgagaacc tcaaggacac tgtagagaa ctaacggcta aa ctggccag caccatagca 2640
gataaccagg agcaagatct ggagaaaaca cggcagtact ct caaaagct agggctgctg 2700
actgagcaac tacagagcct gactctcttt ctacagacaa aa ctaaagga gaagactgaa 2760
caagagacct ttctgctgag tacagcctgt cctcccaccc aggaacaccc totgcctaact 2820
gacaggacct tccctgggaag catcttgaca gcagtggcag at gaagagcc agaactcaact 2880
cctgtgccct tgcttgggaag tgacaagagt gctttcaccc gactagcatc aatgggtttcc 2940
cttcagcccg cagagacccc aggcattggag gagagcctgg cagaaatgag tattatgact 3000
actgagcttc agagtctttg ttccctgcta caagagtcta aa gaagaagc catcaggact 3060
ctgcagcgaa aaatttgtga gctgcaagct aggttgagg cc caggaaga acagcatcag 3120
gaagtccaga aggcaaaaga agcagacata gagaagctga ac caggcctt gtgcttgccg 3180
tacaagaatg aaaaggagct ccaggaagtg atacagcaga atgagaagat cctagaacag 3240
atagacaaga gtggcgagct cataagcctt agagaggagg tg acccacct taccgctca 3300
cttcggcgctg cggagacaga gaccaaagtg tccaggagg cctggcaggc cagctggact 3360
ccaaatgcca gcatatggcc accaatggga ctcaggagaa agtggtggctc tctcaggagg 3420
tgga caaact gagagtgatg ttccctggga tgaaaaatga gaaggaaaac tctgatcaa 3480
gttc cagagc ccatagaaat atcctagagg agaaccctcg gcgctctgac aaggagttag 3540
aaaaactaga tgacattgtt cagcatatct ataagaccct gctctctatt ccagaggtgg 3600
tgaggggatg caaagaacta cagggtatgc tggaaatttct gagctaagaa actgaaagcc 3660
agaaatttgtt tcacctcttt ttacctgcaa taccocctta cc caataacc aagaccaact 3720
ggcatagagc caactgagat aaatgctatt taaataaagt gtattttaatg aaaaaaaaaa 3780
aaaaaaaaa a 3791

```

<210> 208

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006461

<400> 208

ctgaCaagga gttagaaaaa ctagatgaca ttgttcagca tatTTataag accctgctct 60

<210> 209

<211> 2856

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006516

<400> 209

```

tagtcgcggg tccccgagtg agcacgccag ggagcaggag accaaacgac ggggggtcggg 60
gtcagagtcg cagtgggagt ccccgaccg gagcacgagc ctgagcggga gagcgccgt 120
cgcacgcccg tcgccacccg cgtaccggc gcagccagag ccaccagcg agcgtgcga 180
tgagacccag cagcaagaag ctgacgggtc gctcatgtct ggctgtggga ggagcagtc 240
ttgggtccct gcagtttggt tacaacactg gactcatcaa tgccccccag aagggtgatcg 300
aggagttcta caaccagaca tgggtccacc gctatgggga gagcatcctg cccaccacgc 360
tcaccacgct ctggtccttc tcagtggcca tcttttctgt tgggggcatg attgggtcct 420
tctctgtggg ccttttctgt aaccgctttg gccggcggaa ttcaatgctg atgatgaacc 480
tgctggcctt cgtgtccgcc gtgtctcatg gcttctcgaa actgggcaag tcctttgaga 540
tgctgatcct gggccgcttc atcatcgggtg tgtactgcgg cctgaccaca ggcttcgtgc 600
ccatgtatgt ggggaagtgc tcaccacag ccttctctgg ggccctgggc accctgcacc 660
agctgggcat cgtcgtcggc atcctcatcg ccaggtgttt cggcctggac tccatcatgg 720
gcaacaagga cctgtggccc ctgctgctga gcatcatctt catcccgcc ctgctgcagt 780
gcatcgtgct gcccttctgc cccgagagtc cccgcttcct gctcatcaac cgcaacgagg 840
agaaccgggc caagagtgtg ctaaagaagc tgcgcgggac agctgacgtg acccatgacc 900
tgaggagat gaaggaagag agtcggcaga tgatcgggga gaagaaggtc accatcctgg 960
agctgttccg ccccccgcc taccgccagc ccatcctcat cgtgtgggtg ctgcagctgt 1020
cccagcagct gtctggcatc aacgctgtct tctattac tc cagcagcatc ttcgagaagg 1080
cgggggtgca gcagcctgtg tatgccacca ttggctccgg tatcgtcaac acggccttca 1140
ctgtcgtgtc gctgtttgtg gtggagcgag caggccggcg gaccctgcac ctcataggcc 1200
tcgctggcat ggcgggttgt gccatactca tgaccatcgc gctagcactg ctggagcagc 1260
taccctggat gtccatctg agcatcgtgg ccatctttgg ctttgtggcc ttctttgaag 1320
tgggtcctgg ccccatccca tgggtcatcg tggctgaa ct cttcagccag ggtccacgtc 1380
cagctgccat tgccgttgca ggcttctcca actggacc tc aaatttcatt gtgggcatgt 1440
gcttccagta tgtggagcaa ctgtgtggtc cctacgtc tt catcatcttc actgtgctcc 1500
tggttctgtt cttcatcttc acctacttca aagtctctga gactaaaggc cggaccttcg 1560
atgagatcgc ttccggcttc cggcaggggg gagccagc ca aagtgataag acaccgagg 1620
agctgttcca tcccctgggg gctgattccc aagtgtgagt cgcccagat caccagccc 1680
gcctgtccc agcagcccta aggatctctc aggagcac ag gcagctggat gagacttcca 1740
aacctgacag atgtcagccg agccgggccc ggggtctc tt tctccagcca gcaatgatgt 1800
ccagaagaat attcaggact taacggctcc aggtttttat aatttttt ta ttactgattt tgttattttt 1920
aatctatttc agacaagcaa cagggtttat ccaggctt ca ccctgaatgg ttccatgcct 1980
atatcagctc gactctcctg tgcccacatc cagggtct ca caccagcta atctgtaggg 2040
gagggtggag actaagccct gtcgagacac ttgccttc tt caccagcta atctgtaggg 2100
ctggacctat gtcctaagga cacactaatc gaactatgaa ctacaaagct tctatcccag 2160
gaggtggcta tggccacccg ttctgctggc ctggatct cc ccactctagg ggtcaggctc 2220
cattaggatt tgccccttc catctcttcc taccac ca ctcaaattaa tctttcttta 2280
cctgagacca gttgggagca ctggagtgc gggaggag ag gggaagggcc agtctgggct 2340
gccgggttct agtctcctt gcactgaggg ccacacta tt accatgagaa gagggcctgt 2400
gggagcctgc aaactcactg ctcaagaaga catggaga ct cctgccctgt tgtgtataga 2460
tgcaagatat ttatatatat ttttggttgt caatatta aa tacagacact aagttatagt 2520
atatctggac aagccaact gtaaatatc cactcac tc ctgttactta cctaacaga 2580
tataaatggc tggtttttag aaacatgggt ttgaaatg ct tgtggattga gggtaggagg 2640
tttgatggg agtgagacag aagtaagtgg ggttgcaa cc actgcaacgg cttagacttc 2700
gactcaggat ccagtccttc acacgtacct ctcatcag tg tctcttctgt caaaaatctg 2760
tttgatccct gttaccaga gaatatatac attcttta tc ttgacattca aggcatttct 2820
atcacatatt tgatagttgg tgttcaaaaa aacactagt tt ttgtgccagc cgtgatgctc
aggcttga aa tcgcattatt ttgaatgtga agggaa 2856

```

<210> 210

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006516

<400> 210

```

aaacagatat aaatggctgg ttttttagaaa catggttt tg aaatgcttgt ggattgaggg 60

```


<210> 211
 <211> 576
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006607

<400> 211
 atggctactc tgatctacgt tgataaggaa attggagaac caggcaccgc tgtggctgcc 60
 aaggatgtgc tgaagctgga gtctagacct tcaatcaaag cattagatgg gatatactcaa 120
 gttttaacac cactgttttg caaacatac gatgtctccat cagccttacc taaagctacc 180
 agaaaaggctt tgggcactgt caacagagct acagaaaagt cagtaaagac caatggaccc 240
 agaaaaacaaa aacagccaag cttttctgcc aaaaagatga ccgagaagac tgttaaaaca 300
 aaaagtctctg ttcctgcctc agatgacgcc tatccagaaa tagaaaaatt ctttcccttc 360
 aatcttctag actttgagag ttttgacctg cctgaagagc gccagattgc acacctcccc 420
 ttgagtggag tgctctctcat gatccttgat gaggaggag agcttgaaaa gctgtttcag 480
 ctgggcccc cttcacctgt gaaaatgccc tctccacat gggaatgcaa tctgtttgca 540
 gtctccttca agcattctgt cgaccctgga tgttga 576

<210> 212
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006607

<400> 212
 cgcctatcca gaaatagaaa aattctttcc cttcaatctt ctagactttg agagttttga 60

<210> 213
 <211> 2058
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006820

<400> 213
 gcacgaggaa gccacagatc tcttaagaac tttctgtctc caaacctgtg ctgctcgata 60
 aatcagacag aacagttaat cctcaattta agcctgatct aaccctaga aacagatata 120
 gaacaatgga agtgacaaca agattgacat ggaatgatga aaatcatctg cgcaactgtc 180
 tggaaaatgtt tctttgagtc ttctctataa gtctagtgtt catggaggta gcattgaaga 240
 tatggttgaa agatgcagcc gtcagggatg tactataaca atggccttaca ttgattacaa 300
 tatgattgta gcctttatgc ttggaaatta tattaattta cgtgaaagt ctacagagcc 360
 aaatgattcc ctatggtttt cacttcaaaa gaaaaatgrac accactgaaa tagaaacttt 420
 actcttaaat acagcaccaa aaattattga tgagcaactg gtgtgtcgtt tatcgaaaac 480
 ggatattttc attatatgtc gagataataa aatttatcta gataaaatga taacaagaaa 540
 cttgaaacta aggttttatg gccaccgtca gtatttggaa tgtgaagtgt ttcgagttga 600
 aggaattaag gataacctag acgacataaa gaggataa tt aaagccagag agcacagaaa 660
 taggcttcta gcagacatca gagactatag gccctatgca gacttggttt cagaaattcg 720
 tattcttttg gtgggtccag ttgggtctgg aaagtcca gt tttttcaatt cagtcaagtc 780
 tatttttcat ggccatgtga ctggccaagc cgtagtgggg tctgatacca ccagcataac 840
 cgagcgggat aggatatatt ctgttaaaga tggaaaaa at ggaaaatctc tgccatttat 900
 gttgtgtgac actatggggc tagatggggc agaaggagca ggactgtgca tggatgacat 960
 tccccacatc ttaaaaggtt gtatgccaga cagatatcag tttaattccc gtaaaccaat 1020
 tacacctgag cattctactt ttatcacctc tccatctc tg aaggacagga ttcactgtgt 1080
 ggcttatgtc ttagacatca actctattga caatctctac tctaaaatgt tggcaaaagt 1140
 gaagcaagtt cacaagaag tattaaactg tggatatgca tatgtggcct tgcttactaa 1200

```

agtggatgat tgcagtgagg ttcttcaaga caacttttta aacatgagta gatctatgac 1260
ttctcaaagc cgggtcatga atgtccataa aatgctaggc attcctatctt ccaatatctt 1320
gatggttgga aattatgctt cagatttggga actggacccc atgaaggata ttctcatcct 1380
ctctgcaactg aggcagatgc tgcgggctgc agatgatttt ttagaagatt tgcctcttga 1440
ggaaactggt gcaattgaga gagcggttaca gccctgcat ttagataagt tgccttgatt 1500
ctgacatttg gccagcctg tactgggtgtg ccgcaatga agtcaatctc tattgacagc 1560
ctgcttcaga ttttgctttt gttcgttttg ccttctgtcc ttggaacagt catatctcaa 1620
gttcaaaggc caaaacctga gaagcgggtgg gctaagata gtcctactgc aaaccacccc 1680
tccatatctt cgtaccattt acaattcagt ttctgtgaca tctttttaaa ccaactggagg 1740
aaaaatgaga tattctctaa ttattctctt tataacactc tatatagagc tatgtgagta 1800
ctaatacat tgaataatag ttataaaatt attgtataga catctgcttc ttaaacagat 1860
tgtgagttct ttgagaaaca gcgtggattt tacttatctg tgtattcaca gagcttagca 1920
cagtgcctgg taatgagcaa gcatacttgc cattaacttt ccttcccact ctctccaaca 1980
tcacattcac tttaaatttt tctgtatata gaaaggaaa ctagcctggg caacatgatg 2040
aaaccccatc tccactgc 2058

```

<210> 214

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006820

<400> 214

```

tgagttcttt gagaaacagc gtggatttta cttatctgtg tattcacaga gcttagcaca 60

```

<210> 215

<211> 2825

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006845

<400> 215

```

gcgaaattga ggtttcttgg tattgcgcgt ttctcttctt tgctgactct ccgaatggcc 60
atggactcgt cgcttcaggc ccgcctgttt cccgggtctc ctatcaagat ccaacgcagt 120
aatggtttaa ttacacagtgc caatgtaagg actgtgaa ct tggagaaatc ctgtgtttca 180
gtggaatggg cagaaggagg tgccacaaag ggcaagaga ttgattttga tgatgtggct 240
gcaataaacc cagaactctt acagcttctt cctttaca tc cgaaggacaa tctgcccttg 300
caggaaaatg taacaattcca gaaacaaaaa cggagatc cg tcaactccaa aattcctgct 360
ccaaaagaaa gtcttcgaag ccgctccact cgcatgtc ca ctgtctcaga gcttcgcac 420
acggctcagg agaatgacat ggaggtggag ctgcctgc ag ctgcaaactc ccgcaagcag 480
ttttcagttc ctctgcccc cactaggcct tctgccc tg cagtggctga aataccattg 540
aggatgggtc gcgaggagat ggaagagcaa gtccattc ca tccgtggcag ctcttctgca 600
aaccctgtga actcagttcg gaggaatca tgtcttgt ga aggaagtgga aaaaatgaag 660
aacaagcgag aagagaagaa ggcccagaac tctgaaat ga gaatgaagag agctcaggag 720
tatgacagta gttttccaaa ctgggaattt gcccgaaat ga ttaaagaatt tcgggctact 780
ttggaatgtc atccacttac tatgactgat cctatcga ag agcacagaat atgtgtctgt 840
gttaggaaac gccactgaa taagcaagaa ttggccaa ga aagaaattga tgtgatttcc 900
attcctagca agtgtctcct cttggtacat gaacccaa gt tgaaagtgga cttaacaaag 960
tatctggaga accaagcatt ctgctttgac tttgcatt tg atgaaacagc ttcgaaatgaa 1020
gttgtctaca ggttcacagc aaggccactg gtacagac aa tctttgaagg tggaaaagca 1080
acttgttttg catatggcca gacaggaagt ggcaagac ac atactatggg cggagacctc 1140
tctgggaaag ccagaaatgc atccaaaggg atctatgc ca tggcctcccg ggacgtcttc 1200
ctcctgaaga atcaaccctg ctaccggaag ttgggcct gg aagtctatgt gacattcttc 1260
gagatctaca atgggaagct gtttgacctg ctcaacaa ga aggccaaagt gcgcgtgctg 1320
gaggacggca agcaacaggt gcaagtgggt gggctgcagg agcatctggt taactctgct 1380
gatgatgtca tcaagatgct cgacatgggc agcgcctgca gaacctctgg gcagacattt 1440
gccaaactcca attcctcccg ctcccacgcg tgcttccaaa ttattcttct agctaaaggg 1500

```

```

agaatgcatg gcaagttctc tttggtagat ctggcagggga atgagcgagg cgcagacact 1560
tccagtgtctg accggcagac ccgcatggag ggcgcagaaa tcaacaagag tctcttagcc 1620
ctgaaggagt gcatcagggc cctgggacag aacaaggctc acaccccggt ccgtgagagc 1680
aagctgacac aggtgtctgag ggactccttc attggggaga actctaggac ttgcatgatt 1740
gccacgatct caccaggcat aagctcctgt gaatatactt taaacaccct gagatatgca 1800
gacaggggtca aggagctgag cccccacagt gggcccagtg gagagcagtt gattcaaagt 1860
gaaacagaag agatggaagc ctgctctaac agctttaacg aagccatgac tcagatcagg 1920
gaagaggagg aactgtcttc ccagatgtcc agctttaacg aagccatgac tcagatcagg 1980
gagctggagg agaaggctat ggaagagctc aaggagatca tacagcaagg accagactgg 2040
cttgagctct ctgagatgac cgagcagcca gactatgacc tggagacctt tgtgaacaaa 2100
gcggaatctg ctctggccca gcaagccaag catttctcag ccctgcgaga tgtcatcaag 2160
gccttacgcc tggccatgca gctggaagag caggctagca gacaaataag cagcaagaaa 2220
cggccccagt gacgactgca aataaaaaatc tgtttggttt gacaccagc ctcttccttg 2280
gccctcccca gagaactttg ggtacctggt ggggtctaggc aggggtctgag ctgggacagg 2340
ttctggtaaa tgccaagtat gggggcatct gggcccaggg cagctgggga gggggtcaga 2400
gtgacatggg acactccttt tctgttctc agttgtcgcc ctacgagag gaaggagctc 2460
ttagttaccc ttttgtgttg cccttctttc catcaagggg aatgttctca gcatagagct 2520
ttctccgag catcctgcct gcgtggactg gctgctaatt gagagctccc tgggggtgtc 2580
ctggctctgg ggagagagac ggagccttta gtacagctat ctgctggctc taaaccttct 2640
acgcttttg gccgagcact gaatgtcttg tactttaaaa aaatgtttct gagacctctt 2700
tctactttac tgtctcccta ggtcctaga ggatccctac tgttttctgt tttatgtgtt 2760
tatacattgt atgtaacaat aaagagaaaa aataaaaaaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaa 2825

```

<210> 216
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006845

<400> 216
 aaatgtttct gagacctctt tctactttac tgtctcccta ggtcctaga ggatccctac 60

<210> 217
 <211> 823
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007019

<400> 217
 aaacgcgggc gggcggggccc gcagtctctg agttgcagtc gtgttctccg agttcctgtc 60
 tctctgccaa cgccgcccgg atggcttccc aaaaccgcga cccagccgcc actagcgtcg 120
 ccgcccggcg taaaggagct gagccgagcg ggggcgcgcg ccgggggtccg gtgggcaaaa 180
 ggctacagca ggagctgatg accctcatga tgtctggcga taaagggatt tctgccttcc 240
 ctgaatcaga caaccttttc aaatgggtag ggaccatcca tggagcagct ggaacagtat 300
 atgaagacct gaggtataag ctctcgcta agttccccag tggctaccct tacaatgcgc 360
 ccacagtga gttcctcacg ccctgctat accccaacgt ggacaccag ggtaacatat 420
 gcctggacat cctgaaggaa aagtggctcg ccctgtatga tgtcaggacc attctgctct 480
 ccattccagag ccttctagga gaaccaaca ttgatagtc cttgaacaca catgctgccg 540
 agctctggaa aaaccccaca gcttttaaga agtacctgca agaaacctac tcaaagcagg 600
 tcaccagcca ggagcctga cccaggctgc ccagcctgtc cttgtgtcgt ctttttaatt 660
 tttccttaga tggctgtgtc tttttgtgat ttctgtatag gactctttat cttgagctgt 720
 ggtatTTTTTg ttttgttttt gtcttttaaa ttaagcctcg gttgagccct tgtatattaa 780
 ataaatgcat ttttgtcctt ttttagacaa aaaaaaaaaa aaa 823

<210> 218
 <211> 60

<212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007019

<400> 218
 tggaaaaacc ccacagcttt taagaagtac ctgcaagaaa cctactcaaa gcaggtcacc 60

<210> 219
 <211> 2831
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007183

<400> 219
 gaattccgga caggacgtga agatagttgg gtttggaggc ggccgccagg cccaggcccg 60
 gtggacctgc cgccatgcag gacggtaact tcctgctgtc ggccctgcag cctgaggccg 120
 gcgtgtgtct cctggcgctg ccctctgacc tgcagctgga ccgcccgggc gccgaggggc 180
 cggaggccga gcggctgcgg gcagcccgcg tccaggagca ggtccgcgcc cgctcttgc 240
 agctgggaca gcagccgcgg cacaacgggg ccgctgagcc cgagcctgag gccgagactg 300
 ccagaggcac atccaggggg cagtaccaca ccctgcaggc tggcttcagc tctcgctctc 360
 agggcctgag tggggacaag acctcgggct tccggcccat cgccaagccg gcctacagcc 420
 cagcctcctg gtccctccgc tccgcccgtg atctgagctg cagtcggagg ctgagttcag 480
 ccacaaatgg gggcagcgcc tttggggccg ctgggtacgg ggggtgccag cccaccctc 540
 ccatgcccac caggcccgtg tccttccatg agcgcggtgg ggttgggagc cgggcccact 600
 atgacacact ctccctgcgc tcgctgcggc tggggcccgg gggcctggac gaccgctaca 660
 gcctggtgtc tgagcagctg gagcccgccg ccacctccac ctacagggcc tttgcgtacg 720
 agcgccaggc cagctccagc tccagccggg caggggggct ggactggccc gaggccactg 780
 aggtttcccc gagccggacc atccgtgcc cctgcctgag gaccctgcag cgattccaga 840
 gcagccaccg gagccgcggg gtaggcgggg cagtgcgggg ggccgtcctg gagccagtgg 900
 ctcgagcgcc atctgtgcgc agcctcagcc tcagcctggc tgactcgggc cacctgccgg 960
 acgtgcatgg gttcaacagc tacggtagcc accgaacctt gcagagactc agcagcggtt 1020
 ttgatgacat tgacctgcc tcagcagtca agtacctcat ggcttcagac cccaacctgc 1080
 aggtgctggg agcggcctac atccagcaca agtgctacag cgatgcagcc gccaaagaagc 1140
 agggccgcag ccttcaggcc gtgcctaggc tgggtgaagct cttcaaccac gccaaaccagg 1200
 aagtgcagcg ccatgccaca ggtgccatgc gcaacctcat ctacgacaac gctgacaaca 1260
 agctggccct ggtggaggag aacgggattc tcgagctgct gcggacactg cgggagcagg 1320
 atgatgagct tcgcaaaaat gtcacaggga tcctgtggaa cctttcatcc agcgaccacc 1380
 tgaaggaccg cctggccaga gacacgc tgg agcagctcac ggacctggtg ttgagccccc 1440
 tgtcgggggc tgggggtccc cccctcatcc agcagaacgc ctcgaggcg gagatcttct 1500
 acaacgccac cggcttcttc aggaacctca gctcagcctc tcaggccact cgccagaaga 1560
 tgcgggagtg ccacgggctg gtggacgccc tggtcacctc tatcaaccac gccctggacg 1620
 cgggcaaatg cgaggacaag agcgtggaga acgcggtgtg cgtcctgcgg aacctgtcct 1680
 accgcctcta cgacgagatg ccgcgctccg cgctgcagcg gctggagggt cgcgcccgca 1740
 gggacctggc gggggcgccg ccgggagagg tcgtgggctg cttcacgccg cagagccggc 1800
 ggctgcgcga gctgccctc gccgccgatg cgtcacctt cgcgagggtg tccaaggacc 1860
 ccaagggcct cgagtggctg tggagccccc agatcgtggg gctgtacaac cggctgctgc 1920
 agcgtgcga gctcaaccgg cacacgaagg aggcggccgc cggggcgctg cagaacatca 1980
 cggcaggcga ccgcaggtg gcgggggtgc tgagcgcct ggccctggag caggagcgta 2040
 ttctgaacce cctgctagac cgtgtcagga ccgcgcacca ccaccagctg cgctcactga 2100
 ctggcctcat ccgaaacctg tctcggaacg ctaggaaaca ggacgagatg tccaagagg 2160
 tgggtgagcca cctgatcgag aagctgcag gcagcgtggg tgagaagtgc ccccagccg 2220
 aggtgctggg caacatcata gctgtgtca acaacctggg ggtggccagc cccatcgctg 2280
 cccgagacct gctgtatatt gacggactcc gaaagctcat cttcatcaag aagaagcggg 2340
 acagccccga cagtgagaag tcctcccggg cagcatccag ctcctgggcc aacctgtggc 2400
 agtacaacaa gctccaccgt gactttcggg cgaagggcta tcggaaggag gacttcctgg 2460
 gcccataggt gaagccttct ggaggagaag gtgacgtggc ccagcgtcca agggacagac 2520
 tcagctccag gctgcttggc agcccagcct ggaggagaag gctaatacag gaggggcccc 2580

```

tcgctggggc ccctgtgtgc atctttgagg gtcctggggc accaggaggg gcagggtctt 2640
atagctgggg acttggtctt cgcagggcag ggggtggggc agggctcaag gctgctctgg 2700
tgtatggggg ggtgaccag tcacattggc agagggtggg gttggctgtg gcctggcagt 2760
atcttgggat agccagcact gggaataaag atggccatga acagtacaaa aaaaaaaaaa 2820
aaaaggaatt c 2831

```

<210> 220
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007183

<400> 220
 ctggcagtat cttgggatag ccagcactgg gaataaagat ggccatgaac agtcacaaaa 60

<210> 221
 <211> 2815
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007267

<400> 221
 aggaagcggg ggaagggtgaa gtaggaccga attcctgtgc cgaagaggcc tgcagtggga 60
 gagcaggatg ggggtccgg aggtggcgcc caggctctga gctaccctag gtctgcagac 120
 tagcgggcat tggccagaga catggcccag ccactggcct tcatcctcga tgtccctgag 180
 accccagggg accaggggcca gggcccagc ccctatgatg aaagcgaagt gcacgactcc 240
 ttccagcagc tcatccagga gcagagccag tgcacggccc aggaggggct ggagctgcag 300
 cagagagagc gggaggtgac aggaagtagc cagcagacac tctggcggcc cgagggcacc 360
 cagagcacgg ccacactccg catcctggcc agcatgcccc gccgcaccat tggccgcagc 420
 cgaggtgcca tcatctccca gtactacaac cgcacgggtg agcttcgggtg caggagcagc 480
 cggcccctgc tcgggaactt tgtccgctcc gcctggcccc gcctccgcct gtacgacctg 540
 gagctggacc ccacggccct ggaggaggag gagaagcaga gcctcctggt gaaggagttc 600
 cagagcctgg cagtggcaca gcgggaccac atgtctcgcg ggatgccctt aagcctggct 660
 gagaaacgca gcctgcgaga gaagagcagg accccgaggg ggaagtggag gggccagccg 720
 ggcagcggcg gggctctgctc ctgctgtggc cggctcagat atgcctgcgt gctggccttg 780
 cacagcctgg gcctggcgtc gctctccgcc ctgcaggccc tgatgccgtg gcgtacgcc 840
 ctgaagcgca tcggggggcca gttcggctcc agcgtgctct cctacttctt ctttctcaag 900
 accctgctgg ctttcaatgc cctcctgctg ctgctgctgg tggccttcat catgggcoct 960
 caggctgcct tcccaccgc cctgcccggc cctgcccggc tctgcacagg cctggagctc 1020
 ctacagggc cggttgctt caccacacc gtcatgtact acggccacta cagtaacgcc 1080
 acgctgaacc agcgtgtgg cagccccctg gatggcagcc agtgacacc cagggtgggt 1140
 ggctgcctt acaacatgcc cctggcctac ctctccactg tgggcgtgag cttctttatc 1200
 acctgcatca ccctgggtga cagcatggct cactctttcg gggagagcta ccgggtgggc 1260
 agcacctctg gcattccagc catcacctgc ttctgctcct gggactacaa ggtgacgcag 1320
 aagcgggcct ccgcctcca gcaggacaa attcgcaccc ggctgaagga gctgctggcc 1380
 gagtggcagc tgcggcacag cccaggagc gtgtgcggga ggctgcggca ggcggctgtg 1440
 ctggggcttg tgtggtgct gtgtctgggg accgcgtgg gctgcgccgt ggcggtccac 1500
 gtcttctcgg agttcatgat ccagagtcca gaggtgctg gccaggagge tgtgctgctg 1560
 gtccctcccc tggtggttg cctcctcaac ctggggggcc cctacctgtg ccgtgtcctg 1620
 gccgccctgg agcgcctga ctccccggta ctggagggtg acgtggccat ctgcaggaa 1680
 ctcatcctca agctggccat cctggggaca ctgtgctacc actggctggg ccgcagggtg 1740
 ggcgtcctgc agggccagt ctgggaggat tttgtgggcc aggagctgta ccggttctctg 1800
 gtgatggact tcgtcctcat gttgctggac acgctttttg gggaactgg gtggaggatt 1860
 atctccgaga agaagctgaa gaggaggcgg aagccggagt ttgacattgc ccggaatgtc 1920
 ctggagctga tttatgggca gactctgacc tggctggggg tgctcttctc gccctcctc 1980
 cccgccgtgc agatcatcaa gctgctgctc gtcttctatg tcaagaagac cagccttctg 2040
 gccaaactgcc aggcgcgcgc ccggccctgg ctggcctcac acatgagcac cgtcttctc 2100

```

acgctgctct gcttccccgc cttcctgggc gccgctgtct tcctctgcta cgccgtctgg 2160
caggtgaagc cctcgagcac ctgctggccc ttccggacct tggacaccat gtacgaggcc 2220
ggcaggggtgt ggggtgcgcca cctggaggcg gcaggcccca gggctctctg gctgccctgg 2280
gtgcaccggt acctgatgga aaacaccttc tttgtcttcc tgggtgtcagc cctgctgctg 2340
gccgtgatct acctcaacat ccagggtggtg cggggccagc gcaagggtcat ctgacctgctc 2400
aaggagcaga tcagcaatga ggggtgaggac aaaatcttct taatcaacaa gcttcactcc 2460
atctacgaga ggaaggagag ggaaggagag agcagggttg ggacaaccga ggaggctgcg 2520
gcacccctg cctgctcac agatgaacag gatgcctagg gggacggcga tgggcctcac 2580
gggcccgcgc agcacctga gacacactg ttgctccca gtgacctgc tgggacacca 2640
ggacaaggaa gacagtttct cctctcgaaa gccgcagctg cgcctaggct ggagctggaa 2700
gggtgggtga atccggcttg ggcattccca atgaactctg cctgcttg gactctattt 2760
attctgatta aaggggtttt gcaaatggga aaaaaaaaaa aaaaaaaaaa aaaaa 2815

```

<210> 222

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007267

<400> 222

```

ggtgaggaca aaatcttctt aatcaacaag cttcactcca tctacgagag gaaggagagg 60

```

<210> 223

<211> 1893

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007274

<400> 223

```

atttaccgcc gcgcggagag tga.ggggcca agtccgcctt gctccgccac ttaggccgcc 60
ccagacgctt cctcgggggc tgcacccggg tcgggcgcgg ctgccgcggc tagcgggcct 120
tccccgcacc ggcgcggccc aacgccacc gaaccttctg gaagcggcgg ctgcctgggc 180
ccccacgccg ccagaatcgt acgcccgcgc gagctctctg cagccttggc ggcctgggag 240
gcggggctcg ggggtggggc ggcgcggggg cggggtcggc gcggggaggg cgcgttcgat 300
tcgcccccg gcgcaggcc ccgcctcacc agcccatcg ctccacctct gccctcccc 360
tttatggcgc ggcccgggct catctattcc gcgcggggcc tgccagacac ctgcgcctt 420
ctgcagccgc ccgcgcgcat cgcgcgcgca gccccagca tgcggggccc agacgtcgag 480
acgccgtccg ccctccagat ctgcccggat atcgaggagg caggcgccat catcagacc 540
aatgtccacg gggggaccat cctgaagatg atcgaggagg caggcgccat catcagacc 600
cggcattgca acagccagaa cggggagcgc tgtgtggccg ccttggtctg tgtcagcgc 660
accgacttcc tgtctcccat gtgcatcggt gaggtggcgc atgtcagcgc ggagatcacc 720
tacacctcca agcactctgt gga.ggtgcag gtcaacgtga tgtccgaaaa catcctcaca 780
ggtgccaaaa agctgaccaa taa.ggccacc ctgtggtatg tgccctgtc gctgaagaat 840
gtggacaagg tcctcgaggt gcc.tcctgtt gtgtattccc ggcaggagca ggaggaggag 900
ggccggaagc ggtatgaagc cca.gaagctg gagcgcattg agaccaagtg gaggaacggg 960

```

```

gacatcgctc agccagtcct caa.cccagag ccgaacactg tcagctacag ccagtccagc 1020
ttgatccacc tgggtggggc tttagactgc accctgcacg gctttgtgca cggagggtgtg 1080
accatgaagc tcatggatga ggtcgccggg atcgtggctg cagccactg caagaccaac 1140
atcgtcacag cttccgtgga cgc.cattaat tttcatgaca agatcagaaa aggctgcgtc 1200
atcaccatct cgggacgcat gac.cttcacg agcaataagt ccatggagat cgagggtgtg 1260
gtggacgccg accctgttgt gga.cagctct cagaagcgt accgggccc cagtgccttc 1320
ttcacctacg tgtcgctgag cca.ggaaggc aggtcgctgc ctgtgcccc gctggtgccc 1380
gagaccgagg acgagaagaa gcgcttttag gaaggcaaag ggcggtacct gcagatgaag 1440
gcgaagcgac agggccacgc gga.gcctcag ccctagactc cctcctcctg ccactgggtgc 1500
ctcagtagtc catggcaacg ggc.cagtggt ccagtcactt agaagttccc cccttgcca 1560
aaaacccaat tcacattgag agctgggtgt gtctgaagtt ttcgtatcac agtgtaaac 1620

```

tgtactctct	cctgcaaacc	tacacaccaa	agctttat	atatcattcc	agtatcaatg	1680
ctacacagtg	ttgtcccag	cgccgggagg	cggtgggcag	aaaccctcgg	gaatgcttcc	1740
gagcacgctg	taggggtatg	gaagaaccca	gcaccactaa	taaagctgct	gcttggctgg	1800
aaaaaaaaaa	aaaaaaaaaa	aaaaa.aaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1860
aaaaaaaaaa	aaaaaaaaaa	aaaaa.aaaaa	aaa			1893

<210> 224
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007274

<400> 224
 acctacacac caaagcttta tttat. atcat tccagtatca atgctacaca gtgtgtgtccc 60

<210> 225
 <211> 4157
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007315

<400> 225

agcggggcgg	ggcgccagcg	ctgcc.ttttc	tccctgccggg	tagtttcgct	ttcctgcgca	60
gagtctgcgg	aggggctcgg	ctgca.ccggg	gggatcgcgc	ctggcagacc	ccagaccgag	120
cagaggcgac	ccagcgcgct	cggga.gaggc	tgcaccgcgc	cgcccccgcc	tagcccttcc	180
ggatcctgcg	cgcagaaaag	tttca.ttttg	tgtatgccat	cctcgagagc	tgtctagggt	240
aacgttcgca	ctctgtgtat	ataac.ctcga	cagtcttggc	acctaacgtg	ctgtgcgtag	300
ctgctccttt	ggttgaatcc	ccagg.cccct	gttggggcac	aagggtggcag	gatgtctcag	360
tggtacgaac	ttcagcagct	tgact.caaaa	ttcctggagc	agggttcacca	gctttatgat	420
gacagttttc	ccaatggaat	cagac.agtac	ctggcacagt	ggttagaaaa	gcaagactgg	480
gagcacgctg	ccaatgatgt	ttcat.ttgcc	accatccggt	ttcatgacct	cctgtcacag	540
ctggatgatc	aatatagtgc	ctttt.ctttg	gagaataact	tcttgctaca	gcataacata	600
aggaaaagca	agcgtaatct	tcagg.rataat	tttcaggaag	acccaatcca	gatgtctatg	660
atcattttaca	gctgtctgaa	ggaag.aaagg	aaaattctgg	aaaacgcccc	gagattttaat	720
caggctcagt	cggggaatat	tcaga.gcaca	gtgatgttag	acaaacagaa	agagcttgac	780
agtaaagtca	gaaatgtgaa	ggaca.agggt	atgtgtatag	agcatgaaat	caagagcctg	840
gaagatttac	aagatgaata	tgact.tcaaa	tgcaaaacct	tgcagaacag	agaacacgag	900
accaatgggtg	tggcaaagag	tgatc.agaaa	caagaacagc	tgttactcaa	gaagatgtat	960
ttaatgcttg	acaataagag	aaaggaagta	gttcacaaaa	taatagagtt	gctgaatgtc	1020
actgaactta	cccagaatgc	cctga.ttaat	gatgaactag	tggagtgga	gcggagacag	1080
cagagcgcct	gtattggggg	gccgc.ccaat	gcttgcttgg	atcagctgca	gaactgggtc	1140
actatagttg	cggagagtct	gcagc.aagtt	cggcagcagc	ttaaaaagtt	ggaggaattg	1200
gaacagaaat	acacctacga	acatg.accct	atcacaaaaa	acaaacaagt	gttatgggac	1260
cgcaccttca	gtctttttcca	gcagc.tcatt	cagagctcgt	ttgtggtgga	aagacagccc	1320
tgcatgccaa	cgcaccctca	gaggc.cgctg	gtcttgaaga	caggggtcca	gttactgtg	1380
aagttgagac	tgttggtgaa	attgc.aagag	ctgaattata	atttgaaaagt	caaagtctta	1440
tttgataaag	atgtgaatga	gagaa.ataca	gtaaaaggat	ttaggaagtt	caacattttg	1500
ggcagcgaca	caaaagtgat	gaaca.tggag	gagtcaccca	atggcagctc	ggcggctgaa	1560
tttcggcacc	tgcaattgaa	agaac.agaaa	aagtctggca	ccagaacgaa	tgagggtcct	1620
ctcatcggtta	ctgaagagct	tcact.cccct	agttttgaaa	cccaattgtg	ccagcctggg	1680
ttggtaattg	acctcgagac	gacct.ctctg	cccgttgtgg	tgatctccaa	cgtcagccag	1740
ctcccgagcg	gttgggcctc	catcc.ttttg	tacaacatgc	tggtggcgga	accaggaat	1800
ctgtccttct	tcctgactcc	accat.gtgc	cgatgggctc	agctttcaga	agtgtctgag	1860
tggcagtttt	cttctgtcac	caaaa.gagggt	ctcaatgtgg	accagctgaa	catgttggga	1920
gagaagcttc	ttggctcctaa	cgcca.gcccc	gatgggtctca	ttccgtggac	gagggttttgt	1980
aaggaaaata	taaatgataa	aaatt.ttccc	ttctggcttt	ggattgaaag	catcctagaa	2040
ctcattaaaa	aacacctgct	ccctc.tctgg	aatgatgggt	gcatcatggg	cttcatcagc	2100

```

aaggagcgag agcgtgccct gt tgaaggac cagcagccgg ggaccttcct gctgcgggttc 2160
agtgagagct cccgggaagg gg ccatcaca ttcacatggg tggagcggtc ccagaacgga 2220
ggcgaacctg acttccatgc ggt tgaaccc tacacgaaga aagaactttc tgctgttact 2280
ttccctgaca tcattcgcaa tt acaaagtc atggctgctg agaattattc tgagaatccc 2340
ctgaagtatc tgtatccaaa ta ttgacaaa gaccatgcct ttggaaagta ttactccagg 2400
ccaaaggaag caccagagcc aa tggaaact gatggcccta aaggaactgg atatatcaag 2460
actgagttga tttctgtgtc tgaagttcac ccttctagac ttcagaccac agacaacctg 2520
ctcccatgt ctctgagga gt ttgacgag gtgtctcgga tagtgggctc tgtagaattc 2580
gacagtatga tgaacacagt at agagcatg aatttttttc atcttctctg gcgacagttt 2640
tccttctcat ctgtgattcc ct cctgctac tctgttcctt cacatcctgt gtttctaggg 2700
aaatgaaaga aaggccagca aa ttcgctgc aacctgttga tagcaagtga attttctct 2760
aactcagaaa catcagttac tc tgaagggc atcatgcac ttactgaagg taaaattgaa 2820
aggcattctc tgaagagtgg gt ttcacaag tgaaaaacat ccagatacac ccaaagtatc 2880
aggacgagaa tgagggtcct tt gggaaagg agaagttaag caacatctag caaatgttat 2940
gcataaagtc agtgcaccaac tgttataggt tgttgataa atcagtgggt atttaggga 3000
ctgcttgacg taggaacggt aa atttctgt gggagaattc ttacatgttt tctttgcttt 3060
aagtgttaact ggcagttttc ca ttggttta cctgtgaaat agttcaaagc caagtttata 3120
tacaattata tcagtcctct tt caaaggta gccatcatgg atctggtagg gggaaaatgt 3180
gtattttatt acatctttca ca ttggctat ttaaagacaa agacaaattc tgtttcttga 3240
gaagagaata ttagctttac tg tttgttat ggcttaatga cactagctaa tatcaataga 3300
aggatgtaca tttccaaatt ca caagttgt gtttgatac caaagctgaa tacattctgc 3360
tttcatcttg gtcacatata at tattttta cagttctccc aaggaggtta ggctattcac 3420
aaccactcat tcaaaagttg aa attaacca tagatgtaga taaactcaga aatttaattc 3480
atgtttctta aatgggctac tt tgtccttt ttgttattag ggtgggtattt agtctattag 3540
ccacaaaatt gggaaaggag ta gaaaaagc agtaactgac aacttgaata atacaccaga 3600
gataaataga gaatcagatc at ttcaaaac tcatttccta tgtaaactgca ttgagaactg 3660
catatgtttc gctgatatat gt gtttttca catttgcgaa tggttccatt ctctctcctg 3720
tactttttcc agacactttt tt gagtggat gatgtttcgt gaagtatact gtatttttac 3780
ctttttcctt ccttatcact ga cacaaaaa gtagattaag agatgggttt gacaagggtc 3840
ttccctttta catactgctg tc tatgtggc tgtatcttgt ttttccacta ctgctaccac 3900
aactatatta tcatgcaaat gc tgtattct tctttgggtg agataaagat ttcttgagtt 3960
ttgtttttaa attaaagcta aa gtatctgt attgcattaa atataatat cacacagtgc 4020
tttcctggc actgcataca at ctgaggcc tctctctca gtttttatat agatggcgag 4080
aacctaagtt tcagttgatt ttacaattga aatgactaaa aaacaaagaa gacaacatta 4140
aaacaatatt gtttcta 4157

```

<210> 226

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007315

<400> 226

atcagatcat ttcaaaactc at ttcctatg taactgcatt gagaactgca tatgtttcgc 60

<210> 227

<211> 1696

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 227

```

caaaggactt cctagtgggt gt gaaaggca gcggtggcca cagaggcggc ggagagatgg 60
ccttcagcgg ttcccaggct cc ctacctga gtccagctgt ccccttttct gggactattc 120
aaggagggtc ccaggacgga ct tcagatca ctgtcaatgg gaccgttctc agctccagtg 180
gaaccagggt tgctgtgaac tt tcagactg gcttcagtgg aaatgacatt gccttccact 240
tcaaccctcg gtttgaagat ggagggtacg tgggtgtgcaa cacgaggcag aacggaagct 300

```



```

ggggggccga ggagaggaag acacacatgc ctttccagaa ggggatgccc tttgacctct 3 60
gcttctcgtg gcagagctca gatttcaagg tgatgggtgaa cgggatcctc ttcgtgcagt 4 20
acttccaccg cgtgcccttc caccgtgtgg acaccatctc cgtcaatggc tctgtgcagc 4 80
tgtcctacat cagcttccag aacccccgca cagtcctgtt tcagcctgcc ttctccacgg 5 40
tgccgttctc ccagcctgtc tgtttccac ccaggcccag ggggcgcaga caaaaacctc 6 00
ccggcggtgt gcctgccaac ccggctccca ttaccagac agtcatccac acagtgcaga 6 60
gcgccccgtg acagatgttc tctactcccg ccattcccac tatgatgtac cccacccccg 7 20
cctatccgat gcctttcatc accaccatc tgggagggtt gtaccatcc aagtccatcc 7 80
tcctgtcagg cactgtcctg ccagtgctc agaggttcca catcaacctg tgctctggga 8 40
accacatcgc cttccacctg aacccccgtt ttgatgagaa tgctgtgggtc cgcaacaccc 9 00
agatcgacaa ctctcggggg tctgaggagc gaagtctgcc ccgaaaaatg cccttcgtcc 9 60
gtggccagag cttctcagtg tggatcttgt gtgaagctca ctgcctcaag gtggccgtgg 1 020
atggtcagca cctgtttgaa tactaccatc gcctgaggaa cctgcccacc atcaacagac 1 080
tggaagtggg gggcgacatc cagctgacct atgtgcagac ataggcggtt tcctggccct 1 140
ggggccgggg gctgggggtg ggggcagtct gggtcctctc atcatcccca cttcccaggc 1 200
ccagcctttc caaccctgcc tgggatctgg gctttaatgc agaggccatg tccttgtctg 1 260
gtcctgcttc tggctacagc caccctggaa cggagaaggc agctgacggg gattgccttc 1 320
ctcagccgca gcagcacctg gggctccagc tgctggaatc ctaccatccc aggaggcagg 1 380
cacagccagg gagaggggag gagtgggcag tgaagatgaa gccccatgct cagtccctc 1 440
ccatcccca cgcagctcca cccagctccc aagccaccag ctgtctgctc ctggtgggag 1 500
gtggcctcct cagccccctc tctctgacct ttaacctcac tctcaccttg caccgtgcac 1 560
caacccttca cccctcctgg aaagcaggcc tgatggcttc cactggcct ccaccacctg 1 620
accagagtgt tctcttcaga ggactggctc ctttcccagt gtccttaaaa taaagaaatg 1 680
aaaatgcttg ttggca 1696

```

<210> 228

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 228

```

cagaggactg gtcctttcc cagtgtcctt aaaataaaga aatgaaaatg cttgttggca 6 0

```

<210> 229

<211> 6552

<212> DNA

<213> Homo sapiens

<300>

<308> NM_012291

<400> 229

```

atgaggagct tcaaaagagt caactttggg actctgctaa gcagccagaa ggaggctgaa 6 0
gagttgctgc ccgacttgaa ggagttcctg tccaaccctc cagctgggtt tcccagcagc 1 20
cgatctgatg ctgagaggag acaagcttgt gatgccatcc tgagggtctg caaccagcag 1 80
ctgactgcta agctagcttg cctaggcat ctggggagcc tgctggagct ggcagagctg 2 40
gcctgtgatg gctacttagt gtctaccca cagcgtctc cctctacct ggaacgaatt 3 00
ctctttgtct tactgcgaa tgctgtgca caaggaagcc cagaggccac actccgcctt 3 60
gctcagcccc tccatgcctg cttggtgcag tgctctcgcg aggctgctcc ccaggactat 4 20
gaggccgtgg ctgggggcag cttttctctg ctttggagg gggcagaagc cctgttggaa 4 80
cggcgagctg catttgagc tcggctgaag gccttgagct tctagtact cttggaggat 5 40
gaaagtaccc cttgtgaggt tcctcacttt gcttctccaa cagcctgtcg agcggtagct 6 00
gcccatcagc tatttgatgc cagtggccat ggtctaaatg aagcagatgc tgatttccta 6 60
gatgacctgc tctccaggca cgtgatcaga gccttggtgg gtgagagagg gagctcttct 7 20
gggcttcttt ctcccagag ggccctctgc ctcttgagc tcaccttggg aactgcccgt 7 80
cgcttttgct ggagccgcca ccatgacaaa gccatcagcg cagtggagaa ggctcacagt 8 40
tacctaagga acaccaatct agcccctagc cttcagctat gtcagctggg ggttaagctg 9 00

```

ctgcagggttg	gggaggaagg	acc tcaggca	gtggccaagc	ttctgatcaa	ggcatcagct	960
gtcctgagca	agagtatgga	ggc accatca	ccccacttc	gggcattgta	tgagagctgc	1020
cagttcttcc	tttcaggcct	gga acgaggc	accaagaggc	gctatagact	tgatgccatt	1080
ctgagcctct	ttgcttttct	tggagggtac	tgctctcttc	tgcagcagct	gcgggatgat	1140
ggtgtgtatg	ggggctcctc	caa gcaacag	cagtcctttc	ttcagatgta	ctttcaggga	1200
cttcacctct	acactgtggt	ggt ttatgac	tttgcccaag	gctgtcagat	agttgatttg	1260
gctgacctga	cccaactagt	gga cagttgt	aaatctaccg	ttgtctggat	gctggaggcc	1320
ttagagggcc	tgtcgggcca	aga gctgacg	gaccacatgg	ggatgaccgc	ttcttacacc	1380
agtaattttg	cctacagctt	cta tagtcac	aagctctatg	ccgaggcctg	tgccatctct	1440
gagccgctct	gtcagcacct	gggttttggtg	aagccaggca	cttatcccgga	ggtgcctcct	1500
gagaagttgc	acaggtgctt	ccggctacaa	gtagagagtt	tgaagaaact	gggtaaacag	1560
gcccagggct	gcaagatggt	gat tttgtgg	ctggcagccc	tgcaaccctg	tagccctgaa	1620
cacatggctg	agccagtcac	ttt ctgggtt	cgggtcaaga	tggatgcggc	cagggctgga	1680
gacaaggagc	tacagctaaa	gac tctgcga	gacagcctca	gtggctggga	cccggagacc	1740
ctggccctcc	tgctgaggga	gga gctgcag	gcctacaagg	cgggtgcgggc	cgacactgga	1800
caggaacgct	tcaacatcat	ctgtgacctc	ctggagctga	gccccgagga	gacaccagcc	1860
ggggcctggg	cacgagccac	cca cctggta	gaactggctc	aggtgctctg	ctaccacgac	1920
tttacgcagc	agaccaactg	ctc tgctctg	gatgctatcc	gggaagccct	gcagcttctg	1980
gactctgtga	ggcctgaggc	cca ggccaga	gatcagcttc	tggacgataa	agcacaggcc	2040
ttgctgtggc	tttacatctg	tac tctggaa	gccaaaatac	aggaagggtat	cgagcgggat	2100
cggagagccc	aggcccctgg	taa ctgggag	gaatttgaag	tcaatgacct	gaactatgaa	2160
gataaactcc	aggaagatcg	ttt cctatac	agtaacattg	ccttcaacct	ggctgcagat	2220
gctgctcagt	ccaaatgcct	gga ccaagcc	ctggccctgt	ggaaggagct	gcttacaaag	2280
gggcaggccc	cagctgtacg	gtgtctccag	cagacagcag	cctcactgca	gatcctagca	2340
gccctctacc	agctggtggc	aaa gcccattg	caggctctgg	aggtcctcct	gctgctacgg	2400
attgtctctg	agagactgaa	gga ccactcg	aaggcactg	gctcctcctg	ccacatcac	2460
cagctcctcc	tgaccctcgg	ctgtcccagc	tatgccctg	tacacctgga	agagtcagca	2520
tcgagcctga	agcatctcga	tca gactact	gacacatacc	tgctcctttc	cctgacctgt	2580
gatctgcttc	gaagtcaact	cta ctggact	caccagaagg	tgaccaaggg	tgtctctctg	2640
ctgctgtctg	tgcttcggga	tcc tgccctc	cagaagtcc	ccaaggcttg	gtacttgctg	2700
cgtgtccagg	tcctgcagct	ggt ggcagct	taccttagcc	tcccgtaaaa	caacctctca	2760
cactccctgt	gggagcagct	ctgtgcccaa	ggctggcaga	cacctgagat	agctctcata	2820
gactcccata	agctcctccg	aagcatcata	ctcctgtcta	tgggcagtga	cattctctca	2880
actcagaaag	cagctgtgga	gacatcgttt	ttggactatg	gtgaaaatct	ggtacaaaaa	2940
tggcagggttc	tttcagaggt	gct gagctgc	tcagagaagc	tgggtctgcca	cctggggccgc	3000
ctgggtagtg	tgagtgaagc	caa ggccttt	tgcttgaggg	ccctaaaact	tacaacaaag	3060
ctgcagatac	cacgccagtg	tgc cctgttc	ctggtgctga	agggcgagct	ggagctggcc	3120
cgcaatgaca	ttgatctctg	tca gtcggac	ctgcagcagg	ttctgttctt	gcttgagtct	3180
tgcacagagt	ttggtggggt	gac tcagcac	ctggactctg	tgaagaagg	ccacctgcag	3240
aaggggaagc	agcaggccca	ggt cccctgt	cctccacagc	tcccagagga	ggagctcttc	3300
ctaagaggcc	ctgctctaga	gct ggtggcc	actgtggcca	aggagcctgg	ccccatagca	3360
ccttctacaa	actcctcccc	agt cttgaaa	accaagcccc	agcccatacc	caacttctctg	3420
tcccattcac	ccacctgtga	ctg ctcgtc	tgcgccagcc	ctgtcctcac	agcagctctg	3480
ctgcgctggg	tattggtcac	ggcagggggtg	aggctggcca	tggggccacca	agccagggtt	3540
ctggatctgc	tgcaggctcg	gct gaagggc	tgtcctgaag	ccgctgagcg	cctcaccctaa	3600
gctctccaag	cttccttgaa	tca taaaaca	ccccctcct	tgggttccaag	cctcttggat	3660
gagatcttgg	ctcaagcata	cac actgttg	gcaactggag	gcctgaacca	gccatcaaac	3720
gagagcctgc	agaaggttct	aca gtcaggg	ctgaagtttg	tagcagcacg	gataccccac	3780
ctagagccct	ggcgagccag	cct gctcttg	atttgggccc	tcacaaaact	aggtggcctc	3840
agctgctgta	ctaccaact	ttt tgcaagc	tcctggggct	ggcagccacc	attaataaaa	3900
agtgctccctg	gctcagagcc	ctc taagact	cagggccaaa	aacgttctgg	acgagggcgc	3960
caaaagttag	cctctgctcc	cct gcgcctc	aataatacct	ctcagaaagg	tctggaagg	4020
agaggactgc	cctgcacacc	taa accccca	gaccggatca	ggcaagctgg	ccctcatgtc	4080
cccttcacgg	tgtttgagga	agt ctgccct	acagagagca	agcctgaagt	acccagggcc	4140
cccagggtac	aacagagagt	cca gacgcgc	ctcaaggtga	acttcagtga	tgacagtga	4200
ttggaagacc	ctgtctcagc	tga ggcctgg	ctggcagagg	agcctaagag	acggggcact	4260
gcttcccggg	gccgggggcg	agc aaggaag	ggcctgagcc	taaagacgga	tgccgtggtt	4320
gccccaggta	gtgcccctgg	gaa ccctggc	ctgaatggca	ggagccggag	ggccaagaag	4380
gtggcatcaa	gacattgtga	gga gcggcgt	ccccagaggg	ccagtgaacca	ggccaggcct	4440

```

ggccctgaga tcctgaggac ca tccctgag gaagaactga ctgacaaactg gagaaaaatg 4500
agcttttgaga tcttcagggg ct ctgacggg gaagactcag cctcaggtgg gaagactcca 4560
gctccggggcc ctgaggcagc tt ctggagaa tgggagctgc tgaggctgga ttccagcaag 4620
aagaagctgc ccagcccatg cc cagacaag gagagtgaca aggaccttgg tctcggctc 4680
cagctccccct cagccccctg ag ccaactggg ttttctaccc tggactccat ctgtgactcc 4740
ctgagtgttg ctttccgggg ca ttagtcac tgtcctccta gtgggctcta tggccacctc 4800
tgccgcttcc tggccttggt cc tggggcac cgggatcctt atgccactgc tttccttgct 4860
accgagtctg tctccatcac ctgtcgccac cagctgctca cccacctcca cagacagctc 4920
agcaaggccc agaagcaccg ag gatcactt gaaatagcag accagctgca ggggctgagc 4980
cttcaggaga tgcttgaga tgtccccctg gcccgcaccc agcgcctctt ttccttcagg 5040
gcttttggaat ctggccactt ccccccagcct gaaaaggaga gtttccagga ggcgctggct 5100
ctgatcccca gtggggtgac tgtgtgtgtg ttggccctgg ccacctcca gcccggaacc 5160
gtgggcaaca cctcctgct ga cccggctg gaaaaggaca gtccccagc cagtgtgcag 5220
attcccaactg gccagaacaa gcttcatctg cgttcagtc tgaatgagtt tgatgccatc 5280
cagaaggcac agaaagagaa cagcagctgt actgacaagc gagaatgggtg gacagggcgg 5340
ctggcactgg accacaggat ggagggtctc atcgcttccc tagagaagtc tgtgctgggc 5400
tgctggaagg ggctgctgct gc cgtccagt gaggagccc gccctgccc ggaggcctcc 5460
cgcctacagg agctgctaca ggaactgtggc tggaaatatc ctgaccgcac tctgctgaaa 5520
atcatgctca gtgggtgccg tgcctcacc cctcaggaca ttcaggccct ggcctacggg 5580
ctgtgcccga cccagccaga gcgagcccag gagctcctga atgaggcagt aggacgtcta 5640
cagggcctga cagtaccaag caatagccac cttgtcttgg tcttagacaa ggacttgacg 5700
aagctgccgt gggaaagcat gcccagcctc caagcactgc ctgtcaccgc gctgccctcc 5760
ttccgcttcc tactcagcta ctccatcatc aaagagtatg gggcctcgcc agtgctgagt 5820
caaggggtgg atccacgaag ta ccttctat gtctgaacc ctcaataa cctgtcaagc 5880
acagaggagc aatttcgagc caatttcagc agtgaagctg gctggagagg agtggttggg 5940
gaggtgccaa gacctgaaca ggtgcaggaa gccctgacaa agcatgatt gtatatctat 6000
gcagggcatg gggctggtgc ccgcttctt gatgggcagg ctgtcctgcg gctgagctgt 6060
cgggcagtgg cctgctgtt tggctgtagc agtgccggcc tggctgtgca tggaaacctg 6120
gagggggctg gcatcgctgct caagtacatc atggctgggt gcccttgtt tctgggtaat 6180
ctctgggatg tgactgaccg cgacattgac cgctacacgg aagctctgct gcaaggctgg 6240
cttgagcag gccagggggc ccccttctc tactatgtaa accaggccc ccaagctccc 6300
cgactcaagt atcttattgg ggcgtgcacct atagcctatg gcttgctgt ctctctgcgg 6360
taaccccatg gagctgtctt attgatgcta gaagcctcat aactgttcta cctccaaggt 6420
tagatttaat ccttaggata actcttttaa agtgattttc cccagtgtt tatatgaaac 6480
atttcctttt gatttaacct cagtataata aagatacatc atttaaacc tgaaaaaaa 6540
aaaaaaaaaa aa 6552

```

<210> 230
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_012291

<400> 230
 agcctcataa ctgttctacc tccaagggtta gatttaatcc ttaggataac tcttttaaag 60

<210> 231
 <211> 6317
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_013261

<400> 231
 tagtaagaca ggtgccttca gttcactctc agtaaggggc tgggtgctg catgagtgtg 60
 tgctctgtgt cactgtggat tggagttgaa aaagcttgac tggcgtcatt caggagctgg 120

atggcggtggg	acatgtgcaa	ccaggactct	gagtctgtat	ggagtgcacat	cgagtgtgct	180
gctctgggtg	gtgaagacca	gcctctttgc	ccagatcttc	ctgaacttga	tctttctgaa	240
ctagatgtga	acgacttgga	tacagacagc	tttctgggtg	gactcaagtg	gtgcagtgac	300
caatcagaaa	taatatccaa	tcagtacaac	aatgagcctt	caaacatatt	tgagaagata	360
gatgaagaga	atgaggcaaa	cttgctagca	gtcctcacag	agacactaga	cagtctccct	420
gtggatgaag	acggattgcc	ctcatttgat	gcgctgacag	atggagacgt	gaccactgac	480
aatgaggcta	gtccttcctc	catgcctgac	ggcacccttc	caccccagga	ggcagaagag	540
ccgtctctac	ttaagaagct	cttactggca	ccagccaaca	ctcagctaag	ttataatgaa	600
tgcagtggtc	tcagtaccca	gaacccatgca	aatcaccaatc	acaggatcag	aacaaaccct	660
gcaattgtta	agactgagaa	ttcatggagc	aataaagcga	agagtatttg	tcaacagcaa	720
aagccacaaa	gacgtccctg	ctcggagctt	ctcaaataatc	tgaccacaaa	cgatgaccct	780
cctcacacca	aaccacaga	gaacagaaac	agcagcagag	acaaatgcac	ctccaaaaag	840
aagtcaccaca	cacagtcgca	gtcacaacac	ttacaagcca	aaccaacaac	tttatctctt	900
cctctgaccc	cagagtcacc	aaatgacccc	aaggggtccc	catttgagaa	caagactatt	960
gaacgcacct	taagtgtgga	actctctgga	actgcaggcc	taactccacc	caccactcct	1020
cctcataaag	ccaaccaaga	taaccctttt	agggcttctc	caaagctgaa	gtcctcttgc	1080
aagactgtgg	tgccaccacc	atcaaagaag	cccaggtaca	gtgagtcctc	tggtacacaa	1140
ggcaataaact	ccaccaagaa	agggccggag	caatccgagt	tgtatgcaca	actcagcaag	1200
tctctagctc	tactgtgtg	acacgaggaa	aggaagacca	agcggcccag	tctgcggctg	1260
tttggtgacc	atgactattg	ccagtcaatt	aattccaaaa	cagaaatact	cattaatata	1320
tcacaggagc	tccaagactc	tagacaacta	gaaaataaag	atgtctcctc	tgattggcag	1380
gggcagatct	gttcttcacc	agattcagac	cagtgtctacc	tgagagagac	tttgagggca	1440
agcaagcagg	tctctccttg	cagcacaaga	aaacagctcc	aagaccagga	aatccgagcc	1500
gagctgaaca	agcacttcgg	tcattcccagt	caagctgttt	ttgacgaaga	agcagacaag	1560
accggtgaac	tgagggacag	tgatttcagt	aatgaacaat	tctccaaact	acctatgttt	1620

ataaattcag	gactagccat	ggatggcctg	tttcatgaca	gcgaagatga	aagtataaaa	1680
ctgagctacc	cttgggatgg	ca.cgcaatcc	tattcattgt	tcaatgtgtc	tccttcttgt	1740
tcttctttta	actctccatg	tagagattct	gtgtcaccac	ccaaatcctt	attttctcaa	1800
agaccccaaa	ggatgcgctc	tcgttcaagg	tccttttctc	gacacagggtc	gtgttcccga	1860
tcaccatatt	ccaggccaag	atcaaggctc	ccaggcgagta	gatcctcttc	aagatcctgc	1920
tattactatg	agtcaagcca	ctacagacac	cgcacgcacc	gaaattctcc	cttgatgtg	1980
agatcacgtt	caagatcgcc	ctacagccgt	cggcccagggt	atgacagcta	cgaggaatat	2040
cagcacgtga	ggctgaagag	ggaagaatat	cgcagagagt	atgagaagcg	agagtctgag	2100
agggccaagc	aaagggagag	gcagaggcag	aaggcaattg	aagagcgccg	tgtgatttat	2160
gtcggtaaaa	tcagacctga	ca.caacacgg	acagaactga	gggaccgttt	tgaagttttt	2220
ggtgaaattg	aggagtgcac	agtaaatctg	cgggatgatg	gagacagcta	tggtttcatt	2280
acctaccgtt	atacctgtga	tgcttttgct	gctcttgaaa	atggatacac	tttgccgagg	2340
tcaaacgaaa	ctgactttga	gc.tgtacttt	tgtggacgca	agcaattttt	caagtctaac	2400
tatgcagacc	tagattcaaa	ctcagatgac	tttgaccctg	cttccaccaa	gagcaagtat	2460

gactctctgg	atttttgatag	tttactgaaa	gaagctcaga	gaagcttgcg	caggtaacat	2520
gttccttagc	tgaggatgac	agagggatgg	cgaataacctc	atgggacagc	gcgtccttcc	2580
ctaaagacta	ttgcaagtca	ta.cttaggaa	tttctcctac	tttacactct	ctgtacaaaa	2640
acaaaacaaa	acaacaacaa	ta.caacaaga	acaacaacaa	caataacaac	aatgggtttac	2700
atgaacacag	ctgctgaaga	gg.caagagac	agaatgatata	ccagtaagca	catgtttatt	2760
catgggtgtc	agcttttgctt	ttcctggagt	ctcttggtga	tggagtgtgc	gtgtgtgcat	2820
gtatgtgtgt	gtgtatgtat	gtgtgtggtg	tgtgtgcttg	gtttagggga	agtatgtgtg	2880
ggtacatgtg	aggactgggg	gcacctgacc	agaatgcgca	agggcaaacc	atttcaaatg	2940
gcagcagttc	catgaagaca	cgcttaaaac	ctagaacttc	aaaatgttcg	tattctattc	3000
aaaaggaaat	atatatatat	atatatatat	atatatatat	atatataaat	taaaaaggaa	3060
agaaaactaa	caaccaacca	ac.caaccaac	caaccacaaa	ccaccctaaa	atgacagccg	3120
ctgatgtctg	ggcatcagcc	tttgtactct	gttttttttaa	gaaagtgcag	aatcaacttg	3180
aagcaagctt	tctctcataa	cgtaaatgatt	atatgacaat	cctgaagaaa	ccacaggttc	3240
catagaacta	atatcctgtc	tctctctctc	tctctctctc	tctctttttt	ttttcttttt	3300
ccttttgcca	tggaatctgg	gtgggagagg	atactgcggg	caccagaatg	ctaaagtttc	3360
ctaacatctt	gaagtttctg	tagttcatcc	ttaatcctga	cacctatgta	aatgtccaaa	3420
atgttgatct	tccactgcaa	atttcaaaaag	ccttgtcaat	ggtcaagcgt	gcagcttggt	3480
cagcggttct	ttctgaggag	cggacaccgg	gttacattac	taatgagagt	tgggtagaac	3540
tctctgagat	gtgttcagat	agtgttaattg	ctacattctc	tgatgtagtt	aagtattttac	3600

```

agatgttaaa tggagtat tttat tttttatg tatatactat acaacaatgt tcttttttgt 3660
tacagctatg cactgtaaat gcagccttct tttcaaaact gctaaat tttt 3720
gaatattcaa atgtaattat gaggtgaaac aattattgta cactaacata tttagaagct 3780
gaacttactg cttatatata tttgattgta aaaacaaaaa gacagtgtgt gtgtctgttg 3840
agtgaacaaa gagcaaaatg atgctttccg cacatccatc ccttaggtga gcttcaatct 3900
aagcatcttg tcaagaaata tcctagtccc cttaaaggat taaccacttc tgcgatattt 3960
ttccacattt tcttgctgct tgtttttctt tgaagtttta tacactggat ttgttagggg 4020
aatgaaattt tctcatctaa aat tttttcta gaagatatca tgattttatg taaagtctct 4080
caatgggtaa ccattaagaa atgtttttat tttctctatc aacagtagtt ttgaaactag 4140
aagtcaaaaa tcttttttaa atgctgtttt gttttaattt ttgtgatttt aatttgatac 4200
aaaatgctga ggtaataatt atagtatgat ttttacaata attaatgtgt gtctgaagac 4260
tatctttgaa gccagtattt ct tcccttg gcagagtatg acgatggat ttatctgtat 4320
tttttacagt tatgcatcct gtataaatac tgatatttca ttcctttgtt tactaaagag 4380
acatatttat cagttgcaga tagcctattt attataaatt atgagatgat gaaaaataa 4440
aagccagtgg aaat ttttcta cctaggatgc atgacaattg tcaggttgga gtgtaagtgc 4500
ttcatttggg aaattcagct tttgcagaag cagtgtttct actgcacta gcatggcctc 4560
tgacgtgacc atgggtgtgt tcttgatgac attgcttctg ctaaatttaa taaaaacttc 4620
agaaaaacct ccattttgat catcaggatt tcatctgagt gtggagtccc tgggaatggaa 4680
ttcagtaaca tttggagtgt gtattcaagt ttctaaattg agattcgatt actgtttggc 4740
tgacatgact tttctggaag acatgatata cctactactc aattgttctt ttcctttctc 4800
tcgccccaca cgatcttgta agatggattt cccccccagg ccaatgcagc taattttgat 4860
agctgcattc atttatcacc agcatattgt gttctgagtg aatccactgt ttgtcctgtc 4920
ggatgcttgc ttgatttttt ggcttcttat ttctaagtag atagaaagca ataaaaatac 4980
tatgaaatga aagaacttgt tcacaggttc tgcgttacaa cagtaacaca tctttaatcc 5040
gcctaattct tgttgttctg taggttaaat gcaggatttt taactgtgtg aacgccaaac 5100
taaagttttc agtctttctt tctgaatttt gagtatcttc tgtttagtaa taataataaa 5160
aagactatta agagcaataa attattttta agaaatcgag atttagtaaa tcctattatg 5220
tgttcaagga ccacatgtgt tctctatttt gcctttaaat ttttgtgaac caatttttaa 5280
tacattctcc tttttgccct ggattgttga catgagtgga atacttggtt tcttttctta 5340
cttatcaaaa gacagcacta cagatatcat attgaggatt aatttatccc cctaccccc 5400
agcctgacaa atattgttac catgaagata gttttcctca atggacttca aattgcatct 5460
agaattagtg gagcttttgt atcttctgca gacactgtgg gtagcccatc aaaatgtaag 5520
ctgtgctcct ctcat tttttta tttttttttt tttgggagag aatatttcaa atgaacacgt 5580
gcaccccatc atcactggag gcaaat tttca gcatagatct gtaggatttt tagaagaccg 5640
tgggccattg ccttcatgcc gtggtaagta ccacatctac aattttggta accgaactgg 5700
tgcttttagta atgtggattt ttttcttttt taaaagagat gtagcagaat aattcttcca 5760
gtgcaacaaa atcaattttt tgctaaacga ctccgagaac aacagttggg ctgtcaacat 5820
tcaaagcagc agagagggaa ct tttgacta ttgggggtatg atgtttgggt cagttgataa 5880
aaggaaacct tttcatgcct ttagatgtga gcttccagta ggtaatgatt atgtgtcctt 5940
tcttgatggc tgtaatgaga acttcaatca ctgtagtcta agacctgac tatagatgac 6000
ctagaatagc catgtactat aatgtgatga ttctaaattt gtacctatgt gacagacatt 6060
ttcaataatg tgaactgctg atttgatgga gctactttaa gattttagg tgaaagtgt 6120
atactgtttg ttgaactatg ctgaagagg aaagtgagcg attagttgag cccttgccgg 6180
gccttttttc cacctgcaa tttctacatgt attgttgtgg ttttattcat tgtatgaaaa 6240
ttcctgtgat ttttttttaa tgtgcagtac acatcagcct cactgagcta ataaagggaa 6300
acgaatgttt caaatct 6317

```

<210> 232

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 232

ctgtagtcta agacctgac tatagatacc tagaatagcc atgtactata atgtgatgat 60

<210> 233

<211> 3237

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 233

gcgaagtga	gggtggccca	ggtggggcca	ggctgactga	atgtatctcc	tagctatgga	60
ctaaataata	catgggggga	aataaacaag	tattcatgag	ggtgaaaatg	tgaccagca	120
ggaaaattac	aactatcttc	aattgacgtt	gaataggatg	agtcattgga	tttaagtgat	180
ttactgaaga	ttatactact	ggtagataga	agagctaaag	aaagatggat	actatgatgc	240
tgaatgtgcg	gaatctgttt	gagcagcttg	tgcgccgggt	ggagattctc	agtgaaggaa	300
atgaagtcca	atctatccag	ttggcgaaag	actttgagga	tttccgtaaa	aagtggcaga	360
ggactgacca	tgagctgggg	aaatacaagg	atcttttgat	gaaagcagag	actgagcgaa	420
gtgctctgga	tgtaagctg	aagcatgcac	gtaatcaggt	ggatgtagag	atcaaacgga	480
gacagagagc	tgaggctgac	tgcgaaaagc	tggaaacgca	gattcagctg	attcgagaga	540
tgctcatgtg	tgacacatct	ggcagcattc	aactaagcga	ggagcaaaaa	tcagctctgg	600
cttttctcaa	cagaggccaa	ccatccagca	gcaatgctgg	gaacaaaaga	ctatcaacca	660
ttgatgaatc	tggttcatt	ttatcagata	tcagctttga	caagactgat	gaatcactgg	720
attgggactc	ttctttgggt	aagactttca	aactgaagaa	gagagaaaag	aggcgctcta	780
ctagccgaca	gtttgtgat	ggtccccctg	gacctgtaaa	gaaaactcgt	tccattggct	840
ctgcagtaga	ccaggggaat	gaatccatag	ttgcaaaaac	tacagtgact	gttcccaatg	900
atggcgggcc	catcgaagct	gtgtccacta	ttgagactgt	gccatatttg	accaggagcc	960
gaaggaaaa	aggtacttta	caaccttgga	acagtgactc	cacctgaac	agcaggcagc	1020
tggagccaag	aactgagaca	gacagtgtgg	gcacgccaca	gagtaatgga	gggatgcgcc	1080
tgcattgactt	tgttcttaag	acggttatta	aacctgaatc	ctgtgttcca	tgtggaaagc	1140
ggataaaaatt	tggcaaatta	tctctgaagt	gtcgagactg	tcgtgtgggtc	tctcatccag	1200

aatgtcggga	ccgctgtccc	cttccctgca	ttcctaccct	gataggaaca	cctgtcaaga	1260
ttggagaggg	aatgctggca	gactttgtgt	cccagacttc	tccaatgac	ccctccattg	1320
ttgtgcattg	tgtaaatgag	attgagcaaa	gaggtctgac	tgagacaggc	ctgtatagga	1380
tctctggctg	tgaccgcaca	gtaaaagagc	tgaagagaa	attcctcaga	gtgaaaactg	1440
taccctctct	cagcaaagtg	gatgatattc	atgctatctg	tagccttcta	aaagactttc	1500
ttcgaaaacct	caaagaacct	cttctgacct	ttcgccttaa	cagagccttt	atggaagcag	1560
cagaaatcac	agatgaagac	aacagcatag	ctgccatgta	ccaagctggt	ggtgaactgc	1620
cccaggccaa	cagggaacaca	ttagctttcc	tcatgattca	cttgacagaga	gtggctcaga	1680
gtccacatac	taaaatggat	gttgccaatc	tggctaaagt	ctttggccct	acaatagtgg	1740
cccatgctgt	gcccaatcca	gacccagtga	caatgttaca	ggacatcaag	cgtcaaccac	1800
aggtgggttg	gcgcctgctt	tccttgcttc	tggagtattg	gagtcagttc	atgatggtgg	1860
agcaagagaa	cattgacccc	ctacatgtca	ttgaaaactc	aaatgccttt	tcaacaccac	1920
agacaccaga	tattaaagtg	agtttactgg	gacctgtgac	cactcctgaa	catcagcttc	1980
tcaagactcc	ttcatctagt	tcctgtgcac	agagagtcct	ttccaccctc	accaagaaca	2040
ctcctagatt	tggagcaca	agcaagctcg	ccactaacct	aggacgacaa	ggcaactttt	2100
ttgcttctcc	aatgctcaag	tgaagtcaaa	tctgcctgtt	acttcccagc	attgactgac	2160
tataagaaag	gacacatctg	tactctgtct	tgcagcctcc	tgtactcatt	actactttta	2220
gcattctcca	ggcttttact	caagtttaat	tgtgcatgag	ggtttttatta	aaactatata	2280
tatctcccc	tcttctcct	caagtcacat	aatatcagca	ctttgtgctg	gtcattgttg	2340
ggagctttta	gatgagacat	ctttccaggg	gtagaagggt	tagtatggaa	ttggttgtga	2400
ttcttttttg	ggaagggggt	tattgttcc	ttggcttaaa	gccaaatgct	gctcatagaa	2460
tgatctttct	ctagtttcat	ttagaactga	tttccgtgag	acaatgacag	aaaccctacc	2520
tatctgataa	gattagcttg	tctcaggggt	ggaagtggga	gggcagggca	aagaaaggat	2580
tagaccagag	gatttaggat	gcctccttct	aagaaccaga	agttctcatt	ccccattatg	2640
aactgagcta	taatatggag	ctttcataaa	aatgggatgc	attgaggaca	gaactagtga	2700
tgggagtatg	cgtagctttg	atttggtatga	ttaggctctt	aatagtgttg	agtggcacaa	2760
ccttgtaaat	gtgaaagtac	aactcgtatt	tatctctgat	gtgccgctgg	ctgaactttg	2820
ggttcatttg	gggtcaaagc	cagtttttct	tttaaaattg	aattcattct	gatgcttggc	2880

ccccataccc	ccaaccttgt	ccagtggagc	ccaacttcta	aaggtcaata	tatcatcctt	2940
tggcatccca	actaacaata	aagagtaggc	tataagggaa	gattgtcaat	atctttgtgg	3000
aagaaaagct	acagtcaatt	tttcttttga	ctttggatgc	tgaaattttt	cccatggaac	3060
atagccacat	ctagataag	gtgagctttt	tcttctgtta	aaattattct	taatgtctgt	3120
aaaaacgatt	ttcttctgta	gaatgtttga	cttcgtattg	acccttatct	gtaaaacacc	3180

tatttgggat aatatttga aaaaaagtaa atagcttttt caaaatgaaa aaaaaa 3237

<210> 234
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_013277

<400> 234
 ctcattcccc attatgaact gagctataat atggagcttt cataaaaatg ggatgcattg 60

<210> 235
 <211> 1122
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_013409

<400> 235
 gtcctcgc ccgcgcctgc cccaggatg gtccgcgcga ggcaccagcc ggggtgggctt 60
 tgcctcctgc tgctgctgct ctgccagttc atggaggacc gcagtgccca ggctgggaac 120
 tgctggctcc gtcaagcgaa gaacggccgc tgccaggtcc tgtacaagac cgaactgagc 180
 aaggaggagt gctgcagcac cggccggctg agcacctcgt ggaccgagga ggacgtgaat 240
 gacaacacac tcttcaagtg gatgatatttc aacgggggag cccccaactg catccctgt 300
 aaagaaacgt gtgagaacgt ggactgtgga cctgggaaaa aatgccgaat gaacaagaag 360
 aacaaacccc gctgcgtctg cgccccggat tgttccaaca tcacctggaa gggctccagt 420
 tgcgggctgg atgggaaaac ctaccgcaat gaatgtgcac tcctaaaggc aagatgtaaa 480
 gagcagccag aactggaagt ccagtaccaa ggcagatgta aaaagacttg tcgggatgtt 540
 ttctgtccag gcagctccac atgtgtggtg gaccagacca ataatgccta ctgtgtgacc 600
 tgtaatcggg tttgccaga gcctgcttcc tctgagcaat atctctgtgg gaatgatgga 660
 gtcacctact ccagtgcctg ccacctgaga aaggctacct gcctgctggg cagatctatt 720
 ggattagcct atgagggaaa gtgtatcaaa gcaaagtcct gtgaagatat ccagtgcact 780
 ggtgggaaaa aatgtttatg ggatttcaag gttgggagag gccgggtgtt cctctgtgat 840
 gagctgtgcc ctgacagtaa gtcggatgag cctgtctgtg ccagtgacaa tgccacttat 900
 gccagcgagt gtgccatgaa ggaagctgcc tgctcctcag gtgtgctact ggaagtaaag 960
 cactccggat cttgcaactc catttcggaa gacaccgagg aagaggagga agatgaagac 1020
 caggactaca gctttcttat atctttctatt ctagagtggg aaactctcta taagtgttca 1080
 gtgttcacat agcctttgtg caaaaaaaaa aaaaaaaaaa aa 1122

<210> 236
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_013409

<400> 236
 gaagatgaag accaggacta cagctttcct atatcttcta ttctagagtg gtaaactctc 60

<210> 237
 <211> 11389
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014246

<400> 237

atggcgccgc	cgccgcgcgc	cgtgctgccc	gtgctgctgc	tcctggccgc	cgccgcgcgc	60
ctgcccggcga	tggggctgcg	agcggccccc	tgggagccgc	gcgtaccgcg	cgggaccgcg	120
gccttcgccc	tccggcccgg	ctgtacctac	gcggtgggcg	ccgcttgca	gccccgggcg	180
ccgcgggagc	tgctggacgt	gggcccgcgt	ggcggtctgg	caggacgtcg	gcgcgtctcg	240
ggcgccggggc	gcccgcgtcc	gctgcaagtc	cgcttggtgg	cccgcagtgc	cccgcaggcg	300
ctgagccgcg	gcctgcgggc	gcgcacgcac	cttcccggct	gcggagcccg	tgcccggctc	360
tgcggaaccg	gtgcccggct	ctgcccggcg	ctctgcttcc	ccgtcccccg	cggtcgccgc	420
gcgcgcgacg	attcggcgct	cgcagctccg	accaccttac	ccgcctgcgc	ctgcccgcgc	480
cgccccaggc	cccgtgtccc	cggccgtccc	atctgcctgc	cgccggggcg	ctcggtccgc	540
ctgcgctctgc	tgtgcgccct	gcggcgccgc	gctggcgccg	tccgggtggg	actggcgctg	600
gaggccgcga	ccgcggggac	gccctccgcg	tcgccatccc	catcgccgc	cctgccgcgc	660
aacttgcccg	aagcccgggc	ggggccggcg	cgacggggccc	ggcggggcac	gagcggcaga	720
gggagcctga	agtttccgat	gcccactac	caggtggcgt	tgtttgagaa	cgaaccggcg	780
ggcaccctca	tcctccagct	gcacgcgcac	tacaccatcg	agggcgagga	ggagcgctg	840
agctattaca	tggaggggct	gttcgacgag	cgctcccggg	gctacttccg	aatcgactct	900
gccacgggcg	ccgtgagcac	ggacagcgta	ctggaccgcg	agaccaagga	gacgcacgtc	960
ctcaggggtga	aagccgtgga	ctacagtacg	ccgcgcgcgt	cgccaccac	ctacatcact	1020
gtcttgggtca	aagacacca	cgaccacagc	ccggtcttcg	agcagtcgga	gtaccgcgag	1080
cgcgtgcggg	agaacctgga	ggtgggctac	gaggtgctga	ccatccgcgc	cagcgaccgc	1140
gactcgccca	tcaacgccaa	cttgcgcttac	cgcggtgttg	ggggcgcgctg	ggagctcttc	1200
cagctcaacg	agagctctgg	cgtggtgagc	acacggggcg	tgctggaccg	ggaggagggc	1260
gcccagtagc	agctcctggt	ggaggccaac	gaccaggggc	gcaatccggg	cccgtcagct	1320
gccacggcca	ccgtgtacat	cgaggtggag	gacgagaacg	acaactacc	ccagttcagc	1380
gagcagaact	acgtggctca	ggtgcccag	gacgtggggc	tcaacacggc	tgtgtgcga	1440
gtgcaggcca	cggaccggga	ccagggccag	aacgcggcca	ttcactacag	catcctcagc	1500
gggaacgtgg	ccggccagtt	ctacctgcac	tcgctgagcg	ggatcctgga	tgtgatcaac	1560
cccttggaatt	tcgaggatgt	ccagaaatac	tcgctgagca	ttaaggccca	ggatgggggc	1620
cggccccccg	tcataaatcc	ttcaggggtg	gtgtctgtgc	aggtgctgga	tgtcaacgac	1680
aacgagccta	tctttgtgag	cagccccctc	caggccacgg	tgctggagaa	tgtgcccctg	1740
ggctaccctg	tgggtgcacat	tcaggcggtg	gacgcggact	ctggagagaa	cgccccgctg	1800
cactatcgcc	tggtggacac	ggcctccacc	ttctgggggg	gcggcagcgc	tgggcctaat	1860
aatcctgccc	ccacccttga	cttccccttc	cagatccaca	acagctccgg	ttggatcaca	1920
gtgtgtgccc	agctggaccg	cgaggaggtg	gagcactaca	gcttcggggg	ggaggcggtg	1980
gaccacggct	cgcccccca	gagctcctcc	accagcgtgt	ccatcacggg	gctggacgtg	2040
aatgacaacg	acccggtgtt	cacgcagccc	acctacgagc	ttcgtctgaa	tgaggatgcg	2100
gccgtgggga	gcagcgtgct	gaccctgcag	gcccgcgacc	gtgacgcca	cagtgtgatt	2160
acctaccagc	tcacaggcgg	caacaccccg	aaccgctttg	cactcagcag	ccagagaggg	2220
ggcgccctca	tcaccttggt	gctacctctg	gactacaagc	aggagcagca	gtacgtgctg	2280
gcggtgacag	catccgacgg	cacacggctc	cacactgcgc	atgtccta	caacgtcact	2340
gatgccaaca	ccacaggcc	tgtctttcag	agctcccatt	acacagttag	tgtcagttag	2400
gacaggcctg	tgggcacctc	cattgctacc	ctcagtccca	acgatgagga	cacaggagag	2460
aatgcccgca	tcacctacgt	gattcaggac	ccggtgcgcg	agttccgcct	tgaccccgac	2520
agtggcacca	tgtacacca	gatggagctg	gactatgaga	accaggctgc	ctacacgctg	2580
accatcatgg	cccaggacaa	cggcatcccg	cagaaatcag	acaccaccac	cctagagatc	2640
ctcatcctcg	atgccaatga	caatgcaccc	cagttcctgt	gggatttcta	ccagggttcc	2700
atctttgagg	atgctccacc	ctcgaccagc	atcctccagg	tctctgccac	ggaccgggac	2760
tcaggctccc	atgggcgtct	gctgtacacc	ttccagggtg	gggacgacgg	cgatggggac	2820
ttctacatcg	agccacgtct	cgggtgtgatt	cgcaccacgc	gcccggctgga	ccggggagaa	2880
gtggccggtg	acaacctttg	ggctctggct	gtggatccgg	gcagtcccac	ttcccttagc	2940
gcctcggtag	aaatccagggt	gaccatcttg	gacattaatg	acaatgccc	catgttttag	3000
aaggacgaac	tggagctgtt	tgttgaggag	aacaaccag	tggggtcggg	gggtggcaag	3060
attcgtgcta	acgaccctga	tgaaggccct	aatgcccaga	tcattgtatca	gattgtggaa	3120
ggggacatgc	ggcatttctt	ccagctggac	ctgctcaacg	gggacctgcg	tgccatgggtg	3180
gagctgggact	ttgagggtccg	gcgggagtat	gtgctgggtg	tgaggccac	gtcggtcccg	3240
ctgggtgagcc	gagccacgggt	gcacatcctt	ctcgtggacc	agaatgacaa	ccgcctgtg	3300
ctgcccgcact	tccagatcct	cttcaacaac	tatgtcacca	acaagtccaa	cagtttccc	3360
accggcggtga	tcggctgcac	cccggcccat	gaccccgacg	tgtcagacag	cctcaactac	3420
accttcgtgc	agggcaacga	gctgcgctg	ttgctgctgg	accccgccac	gggcgaactg	3480
cagctcagcc	gcgacctgga	caacaaccgg	ccgctggagg	cgctcatgga	ggtgtctgtg	3540

tctgatggca	tccacagcgt	cacggccttc	tgcaccctgc	gtgtcaccat	catcacggac	3600
gacatgctga	ccaacagcat	cactgtccgc	ctggagaaca	tgtcccagga	gaagt tccctg	3660
tccccgctgc	tggccctctt	cgtggagggg	gtggccgcgc	tgtgtgccac	caccaaggac	3720
gacgtcttcg	tcttcaacgt	ccagaacgac	accgacgtca	gtcccaacat	cctgaacgtg	3780
accttctcgg	cgctgctgc	tggcggcgct	cgcggccagt	tcttcccgtc	ggaggacctg	3840
caggagcaga	tctacctgaa	tccgacgctg	ctgaccacca	tctccacgca	gcgctgtgctg	3900
cccttcgacg	acaacatctg	cctgcgcgag	ccctgcgaga	actacatgaa	gtgctgtgtcc	3960
gttctgcgat	tgcacagctc	cgcgcccttc	ctcagctcca	ccaccgtgct	cttccggccc	4020
atccacccca	tcaacggcc	gcgctgcgc	tgcgcgccg	gcttcaccgg	cgactactgc	4080
gagacggaga	tgcacctctg	ctactccgac	ccgtgcggcg	ccaacggccg	ctgccgcagc	4140
cgcgagggcg	gctacacctg	cgagtgtctc	gaggacttca	ctggagagca	ctgtgagggtg	4200
gatgcccgt	caggccgctg	tgccaacggg	gtgtgcaaga	acgggggcac	ctgcgtgaac	4260
ctgctcatcg	gcggcttcca	ctgcgtgtgt	cctcctggcg	agtatgagag	gccctactgt	4320
gaggtgacca	ccaggagctt	cccgcgccag	tcttcgtca	ccttcggggg	cctgagacag	4380
cgcttccact	tcaccatctc	cctcacgttt	gccactcagg	aaaggaacgg	cttgc tttctc	4440
tacaacggcc	gcttcaatga	gaagcacgac	ttcatcgccc	tggagatcgt	ggacgagcag	4500
gtgcagctca	ccttctctgc	aggcgagaca	acaacgaccg	tggcaccgaa	ggttc cagct	4560
ggtgtgagtg	acgggcggtg	gcactctgtg	caggtgcagt	actacaacaa	gccccatatt	4620
ggccacctgg	gcttgcctca	tgggcccgtc	ggggaaaaga	tggccgtggt	gacagtggat	4680
gattgtgaca	caaccatggc	tgtgcgcttt	ggaaaggaca	tcgggaacta	cagctgcgct	4740
gcccagggca	ctcagaccgg	ctccaagaag	tccttgatc	tgaccggccc	tctactcctg	4800
gggggtgtcc	ccaacctgc	agaagacttc	ccagtgcaca	accggcagtt	cgtgggctgc	4860
atgcgggaacc	tgtcagtcga	cggcaaaaat	gtggacatgg	ccggaattcat	cgccaacaat	4920
ggcaccgggg	aaggctgcgc	tgtctcgagg	aacttctgcg	atgggaggcg	gtgtcagaat	4980
ggaggcacct	gtgtcaacag	gtggaatatg	tatctgtgtg	agtgtccact	ccgat tccggc	5040
gggaagaact	gtgagcaagc	catgcctcac	cccagctctc	tcagcgggtga	gagcgtcctg	5100
tcctggagtg	acctgaaca	catcatctct	gtgccctggt	acctggggct	catgt tccgg	5160
acccggaagg	aggacagcgt	tctgatggag	gccaccagtg	gtggggccac	cagct ttcgc	5220
ctccagatcc	tgaacaacta	cctccagttt	gaggtgtccc	acggcccctc	cgatgtggag	5280
tccgtgatgc	tgtccgggtt	gcgggtgacc	gacggggagt	ggcaccacct	getga tccgag	5340
ctgaagaatg	ttaaggagga	cagtgcagatg	aagcacctgg	tcaccatgac	cttggactat	5400
gggatggacc	agaacaaggc	agatatcggg	ggcatgcttc	ccgggctgac	ggtaaggagc	5460
gtggtggctg	gaggcgctc	tgaagacaag	gtctccgtgc	gccgtggatt	ccgaggctgc	5520
atgcagggag	tgaggatggg	ggggacgccc	accaacgtcg	ccacctgaa	catgaacaac	5580
gcactcaagg	tcagggtgaa	ggacggctga	gatgtggacg	accctgtac	ctcagcccc	5640
tgtcccccca	atagccgtcg	ccacgacgct	tgggagagct	acagctgcgt	ctgtgacaaa	5700
gggtaccttg	gaataaactg	tgtggatgcc	tgtcacctga	accctgcga	gaacatgggg	5760
gcctgcgtgc	gctccccggg	ctccccgcag	ggctacgtgt	gcgagtgtgg	gccccgtcac	5820
tacgggcegt	actgtgagaa	caaactcgac	cttccgtgcc	ccagaggctg	gtgggggaac	5880
cccgtctgtg	gacctgccca	ctgtgccgtc	agcaaaggct	ttgatcccga	ctgtaataag	5940
accaacggcc	agtgccaatg	caaggagaat	tactacaagc	tcctagccca	ggacaacctgt	6000
ctgccttgcg	actgcttccc	ccatggctcc	cacagccgca	cttgcgacat	ggccaacggg	6060
cagtgtgcct	gcaagcccgg	cgtcatcggc	cgccagtga	accgtgcga	caaccctgtt	6120
gccgaggtca	ccacgctcgg	ctgtgaagtg	atctacaatg	gctgtcccaa	agcat ttag	6180
gccggcatct	ggtggccaca	gaccaagttc	gggcagccgg	ctgcggtgcc	atgcc ttaag	6240
ggatccgttg	gaaatgcggt	ccgacactgc	agcggggaga	agggctggct	gccccagag	6300
ctctttaact	gtaccacca	ctccttcgtg	gacctcaggg	ccatgaatga	gaagctgagc	6360
cgcaatgaga	cgcaggtgga	cggcgccagg	gcctgcagc	tggtgagggc	gctgcgcagt	6420
gctacacagc	acacgggca	gctctttggc	aatgacgtgc	gcacggccta	ccagctgctg	6480
ggccacgtcc	ttcagcacga	gagctggcag	cagggcttcg	acctggcagc	cacgcaggac	6540
gccgactttc	acgaggacgt	catccactcg	ggcagcgccc	tcttgcccc	agccaacagg	6600
gcggcgctggg	agcagatcca	gcggagcgag	ggcggcacgg	cacagctgct	ccggcgcttc	6660
gagggctact	tcagcaacgt	ggcacgcaac	gtgcggcgga	cgtacctgcg	gcccttcgtc	6720
atcgtcaccg	ccaacatgat	tcttgcgtgc	gacatctttg	acaagttcaa	ctttaaggga	6780
gccagggctc	cgcgattcga	caccatccat	gaagagttcc	ccagggagct	ggagtcttcc	6840
gtctccttcc	cagccgactt	cttcagacca	cctgaagaaa	aagaaggccc	cctgctgagg	6900
ccggctggcc	ggaggacca	ccgcagacc	acgcgcccgg	ggcctggcac	cgagagggag	6960
gccccgatca	gcaggcgga	gcgacaccct	gatgacgctg	gccagttcgc	cgtgcctctg	7020
gtcatcattt	accgcacctt	ggggcagctc	ctgcccagac	gctacgaccc	cgaccgtcgc	7080
agcctccggg	tgcctcaccg	gcccatcatt	aataacccga	tggtgagcac	gctgggtgtac	7140
agcgaggggg	ctccgctccc	gagaccctgc	gagaggcccc	tcctgggtgga	gttcgcctctg	7200

ctggaggtgg	aggagcgaac	caagcctgtc	tgcgtgttct	ggaaccactc	cctggc cgtt	7260
ggtgggacgg	gaggggtggtc	tgcccggggc	tgcgagctcc	tgtccaggaa	ccggacacat	7320
gtcgctgcc	agtgcagcca	cacagccagc	tttgcggtgc	tcatggatat	ctccaggcgt	7380
gagaacgggg	aggtcctgcc	tctgaagatt	gtcaccctatg	ccgctgtgtc	cttgtcactg	7440
gcagccctgc	tggtagccctt	cgtcctcctg	agcctgggtcc	gcatgtgtcg	ctccaa cctg	7500
cacagcattc	acaagca cct	cgcctgtggc	ctcttctctc	ctcagctggt	gttctg gatt	7560
gggatcaacc	agacggaaaa	cccgtttctg	tgcacagtgg	ttgccatcct	cctcca ctac	7620
atctacatga	gcacctttgc	ctggaccctc	gtggagagcc	tgcattgtcta	ccgcat gctg	7680
accgaggtgc	gcaacatcga	cacggggccc	atgcggttct	actacgtcgt	gggctg gggc	7740
atcccgga	ttgtcacagg	actggcggtc	ggcctggacc	cccagggtta	cgggaa cccc	7800
gactttctgt	ggctgtcgct	tcaagacacc	ctgattttgga	gctttg cggg	gcccac cgga	7860
gctgttataa	tcatcaaacac	agtcacttct	gtcctatctg	caaaggtttc	ctgcca aaga	7920
aagcaccatt	attatgggaa	aaaagggtatc	gtctccctgc	tgaggaccgc	attcct cctg	7980
ctgctgtctca	tcagcgccac	ctggctgtctg	gggctgtctg	ctgtgaaccg	cgatgc actg	8040
agcttttct	acctcttctgc	catcttcagc	ggcttacagg	gccccttctg	cctcct tttc	8100
cactgcgtgc	tcaaccaggga	ggtccggaag	cacctgaagg	gcgtgtctcg	cgggag gaag	8160
ctgcacctgg	aggactccgc	caccaccagg	gccacctgc	tgacgcgtc	cctcaa ctgc	8220
aacaccacct	tcggtgacgg	gcctgacatg	ctgcgcacag	acttgggcga	gtccac cgcc	8280
tcgctggaca	gcacgtcag	ggatgaagg	atccagaagc	tcggcgtgtc	ctctgg gctg	8340
gtgaggggca	gccacggaga	gccagacgcg	tccctcatgc	ccaggagctg	caagga tccc	8400
cctggccacg	attccgactc	agatagcgag	ctgtccctgg	atgagcagag	cagctc ttac	8460
gcctcctcac	actcgtcaga	cagcgaggac	gatgggggtg	gagctgagga	aaaatg ggac	8520
cgggccagg	gcgcgctca	cagcaccccc	aaaggggacg	ctgtggccaa	ccacgt tccg	8580
gccggctggc	ccgaccagag	cctggctgag	agtgcacagt	aggaccccag	cggcaa gccc	8640
cgctgaagg	tggagacc aa	ggtcagcgtg	gagctgcacc	gcgaggagca	gggcagt cacc	8700
cgtggagagt	accccccgga	ccaggagagc	gggggcgcag	ccaggcttgc	tagcag ccag	8760
ccccagagc	agaggaaa gg	catcttgaaa	aataaagtca	cctaccgcgc	gccgct gacg	8820
ctgacggagc	agacgctgaa	gggcccgtc	cgggagaagc	tggccgactg	tgagca gacg	8880
cccacatcct	cgcgacgctc	tccctgggc	tctggcgcc	ccgactgcgc	catcac agtc	8940
aagagccctg	ggagggagcc	ggggcgtgac	cacctcaacg	gggtggccat	gaatgt gcgc	9000
actgggagcg	cccaggccga	tggctccgac	tctgagaaac	cgtgaggcaa	gcccgt cacc	9060
ccacacaggc	tgcggcat ca	ccctcagacc	ttggagccca	aggggccact	gccctt gaag	9120
tggagtgggc	ccagagtgtg	gcgggtccca	tggtagccca	cccccgactg	atcatc caga	9180
cacaaaggtc	ttggttct cc	caggagctca	gggctgtca	gacctggtga	caagtgc caa	9240
aggccacagg	catgagggag	gcgtggacca	ctggggccagc	accgctgagt	cctaag actg	9300
cagtcaaagc	cagaactgag	aggggacccc	agactgggccc	cagaggctgg	ccagagt tca	9360
ggaacgccgg	gcacagac ca	aagaccgcgg	tccagccccg	cccaggcggg	catctcatgg	9420
cagtgcggac	ccgtggctgg	cagcccgggc	agtcctttgc	aaaggcacc	cttgtct taa	9480
aatcacttctg	ctatgtggga	aaggtggaga	tactttttata	tatttgtatg	ggactct gag	9540
gaggtgcaac	ctgtatat at	attgcattctg	tgctgacttt	gttatcccga	gagatccatg	9600
caatgatctc	ttgtgtctt	ctctgtcaag	attgcacagt	tgtacttgaa	tctggcatgt	9660
gttgacgaaa	ctgggtgcc cc	agcagatcaa	aggtgggaaa	tacgtcagca	gtgggg ctaa	9720
aaaccaagcgg	ctagaagccc	tacagctgcc	ttcggccagg	aagtgaggat	ggtgtgggccc	9780
ctccccgccg	gccccctggg	tcccagtggt	tcgctgtgtg	tcgctttgtc	ctctgctgcc	9840
atctgccccg	gctgtgtgaa	ttcaagacag	ggcagtgcag	cactaggcag	gtgtgaggag	9900
ccctgctgag	gtcactgtgg	ggcacgggtg	ccacacggct	gtcatttttc	acctggtcat	9960
tctgtgacca	ccaccccc tc	ccctcaccgc	ctcccagggtg	gcccggggagc	tgcagg tggg	10020
gatggctttg	tcccttgc tc	ctgctccccg	tgggacctgg	gaccttaaag	cgttgca ggt	10080
tccctgatttg	gacagagg tg	tggggcccttc	caggccgtta	catacctcct	gccaat tctc	10140
taactctctg	agactgcgag	gatctccagg	cagggttctc	ccctctggag	tctgacc aat	10200
tacttcatatt	tgcttcaaat	ggccaattgt	gcagagggac	aaagccacag	ccacact ctt	10260
caacggttac	caaactgttt	ttggaaatcc	acaccaagggt	cgggcccact	gcaggc agct	10320
ggcacagcgt	ggcccagagg	gctgtggaac	gggtcccggga	actgtcagac	atgtttgatt	10380
ttagcgtttc	ctttgttctt	caaatacagg	gccc aaataa	gtgatcagca	cagctgc ttc	10440
caaataaggag	aaaccataaa	ataggatgaa	aatcaagtaa	aatgcaaaga	tgtccac act	10500
gttttaaaact	tgaccctgat	gaaaatgtga	gcactgttag	cagatgccta	tgggaga gga	10560
aaagcgtatc	tgaaaatgggt	ccaggacagg	aggatgaaat	gagatcccag	agtcct caca	10620
cctgaatgaa	ttatacatgt	gccttaccag	gtgagtggctc	tttcgaagat	aaaaaac tct	10680
agtcoccttta	aacgtttgccc	cctggcggtt	cctaagtagc	aaaagggtttt	taagtct tctg	10740
aacagtctcc	tttcatgact	ttaacaggat	tctgccccct	gaggtgtaat	ttttttgttc	10800

tat	ttttttttc	cacgtactcc	acagccaaca	tcacgaggtg	taatttttaa	tttgatcaga	10860
act	gtttacca	aaaaacaact	gtcagtttta	ttgagatggg	aaaaatgtaa	acctatcttt	10920
att	actttaag	actttatggg	agagattaga	cactggaggt	ttttaacaga	acgtgtatct	10980
atta	atgttc	aaaacactgg	aattacaaat	gagaagagtc	tacaataaat	taagatcttt	11040
gaatt	gttac	ttctgcggtg	ctggtttttc	tccacaaaca	cccccgcccc	tccccatgcc	11100
cagggt	ggcc	gtggaaggga	cggtttacgg	acgtgcagct	gagctgtccg	tgtcccattgc	11160
tccct	cagcc	agtggaacgt	gccggaactt	tttgtccatt	ccctagtagg	cctgccacag	11220
cctagat	ggg	cagtttttct	ctttcaccaa	atttgaggac	tttttttttt	tgccatctatt	11280
tcttcag	ttt	tcttttcttg	cactgatctt	tctcctctcc	ttctgtgact	ccagtgtactc	11340
agacgt	taga	cctcttgatg	ttttcccact	ggtccctgag	gctctgttc		11389

<210> 238

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 238

gggagagatt	agacactgga	ggttttttaac	agaacgtgta	tttattaatg	ttcaaaacac	60
------------	------------	-------------	------------	------------	------------	----

<210> 239

<211> 4372

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 239

tagttattaa	agttcctatg	cagctccgcc	tgcggtccgg	cctcatttcc	tccgaaaatc	60
cctgctttcc	ccgctcgcca	cgccctcctc	ctaccgggct	ttaaagctag	tgaggcacag	120
cctgcgggga	acgtagctag	ctgcaagcag	aggccggcat	gaccaccgag	cagcga cgca	180
gcctgcaagg	cttccaggat	tatatccgga	agaccctgga	ccctacctac	atctcragct	240
acatggcccc	ctggtttagg	gaggaagagg	tgcagtatat	tcaggctgag	aaaaac aaca	300
agggcccaat	ggaggctgcc	acactttttc	tcaagttcct	gttggagctc	caggag gaag	360
gctgggtccg	tggctttttg	gatgccctag	accatgcagg	ttattctgga	ctttat gaag	420
ccattgaaag	ttgggatttc	aaaaaaattg	aaaagttgga	ggagtataga	ttactt ttaa	480
aacgtttaca	accagaattt	aaaaccagaa	ttatcccaac	cgatatcatt	tctgat ctgt	540
ctgaatgttt	aattaatcag	gaatgtgaag	aaattctaca	gatttgcctc	actaagg gga	600
tgatggcagg	tgcagagaaa	ttgggtggaat	gccttctcag	atcagacaag	gaaaac tggc	660
ccaaaacttt	gaaacttgc	ttggagaaaag	aaaggaacaa	gttcagtga	ctgtggattg	720
tagagaaaagg	tataaaagat	gttgaaacag	aagatcttga	ggataagatg	gaaact tctg	780
acatacagat	tttctacca	gaagatccag	aatgccagaa	tcttagtgag	aattca tgtc	840
caccttcaga	agtgtctgat	acaaacttgt	acagcccatt	taaaccaaga	aattac caat	900
tagagcttgc	tttgcctgct	atgaaaggaa	aaaacacaat	aatatgtgct	cctacagggt	960
gtggaaaaaac	ctttgtttca	ctgcttatat	gtgaacatca	tcttaaaaaa	ttccca caag	1020
gacaaaaggg	gaaagttgtc	tttttttgca	atcagatccc	agtgtatgaa	cagcagaaat	1080
ctgtattctc	aaaatacttt	gaaagacatg	ggtatagagt	tacaggcatt	tctggag gca	1140
cagctgagaa	tgtcccagtg	gaacagattg	ttgagaacaa	tgacatcatc	atttta actc	1200
cacagattct	tgtgaacaa	cttaaaaagg	gaacgattcc	atcactatcc	atctttactt	1260
tgatgatatt	tgatgaatgc	cacaacacta	gtaaacaaca	cccgtacaat	atgatcatgt	1320
ttaattatct	agatcagaaa	cttggaggat	cttcaggccc	actgccccag	gtcattgggc	1380
tgactgcctc	ggttggtgtt	ggggatgcca	aaaacacaga	tgaagccttg	gattatatct	1440
gcaagctgtg	tgtctctctt	gatgcgtcag	tgatagcaac	agtcaaacac	aatctggagg	1500
aactggagca	agttgtttat	aagccccaga	agtttttcag	gaaagtggaa	tcacggatta	1560
gcgacaaatt	taaatacatc	atagctcagc	tgatgagggg	cacagagagt	ctggca aaga	1620
gaatctgcaa	agacctcgaa	aacttatctc	aaattcaaaa	taggggaattt	ggaaca caga	1680
aatatgaaca	atggatttgt	acagttcaga	aagcatgcat	ggtgttccag	atgccagaca	1740
aagatgaaga	gagcaggat t	tgtaaagccc	tgtttttata	cacttcacat	ttgcggaaat	1800

ataatgatgc	cctcattatc	agtgagcatg	cacgaatgaa	agatgctctg	gattacttga	1860
aagacttctt	cagcaatgtc	cgagcagcag	gattcgatga	gattgagcaa	gatcttactc	1920
agagatttga	agaaaagctg	caggaactag	aaagtgtttc	cagggatccc	agcaatgaga	1980
atcctaaact	tgaagacctc	tgccttcatct	tacaagaaga	gtaccactta	aaccagaga	2040
caataacaat	tctctttgtg	aaaaccagag	cacttggtga	cgctttaaaa	aattggattg	2100
aaggaaatcc	taaactcagt	tttctaaaac	ctggcatatt	gactggacgt	ggcaaaacaa	2160
atcagaacac	aggaatgacc	ctcccggcac	agaagtgtat	attggatgca	ttcaaagcca	2220
gtggagatca	caatattctg	attgccacct	cagttgctga	tgaaggcatt	gacattgcac	2280
agtgcattct	tgtcatcctt	tatgagtatg	tgggcaatgt	catcaaaatg	atccaaacca	2340
gaggcagagg	aagagcaaga	ggtagcaagt	gcttccttct	gactagtaat	gctgggtgta	2400
ttgaaaaaga	acaaataaac	atgtacaaag	aaaaaatgat	gaatgactct	attttacgcc	2460
ttcagacatg	ggacgaagca	gtatttaggg	aaaagattct	gcatatacag	actcatgaaa	2520
aattcatcag	agatagtcaa	gaaaaaccaa	aacctgtacc	tgataaggaa	aataaaaaac	2580
tgctctgcag	aaagtgc aaa	gccttggcat	gttacacagc	tgacgtaaga	gtgatagagg	2640
aatgccatta	cactgtgctt	ggagatgctt	ttaagggaatg	ctttgtgagt	agaccacatc	2700
ccaagccaaa	gcagttttca	agttttgaaa	aaagagcaaa	gatattctgt	gcccagacaga	2760
actgcagcca	tgactgggga	atccatgtga	agtacaagac	atttgagatt	ccagttataa	2820
aaattgaaag	ttttgtggtg	gaggatattg	caactggagt	tcagacactg	tactcgaagt	2880
ggaaggactt	tcatttttgag	aagataccat	ttgatccagc	agaaatgtcc	aaatgatatc	2940
aggtcctcaa	tcttcagcta	caggggaatga	gtaactttga	gtggagaaga	aacaaacata	3000
gtgggtataa	tcatggatcg	cttgtacccc	tgtgaaaata	tattttttaa	aaatatcttt	3060
agcagtttgt	actatattat	atatgcaaag	cacaaatgag	tgaatcacag	cactgagtat	3120
tttgtaggcc	aacagagctc	atagtacttg	ggaaaaatta	aaaagcctca	tttctagcct	3180
tctttttaga	gtcaactgcc	aacaaacaca	cagtaatcac	tctgtacaca	ctgggataga	3240
tgaatgaatg	gaatgttggg	aatttttatc	tccctttgtc	tccttaacct	actgtaacct	3300
ggcttttgcc	cttaacaatc	tactgaaatt	gttcttttga	aggttaccag	tgactctggt	3360
tgccaaatcc	actgggcact	tcttaacctt	ctatttgacc	tctgcgcatt	tggccctggt	3420
gagcactctt	cttgaagctc	tccctgggct	tctctctctt	ctagttctat	tctagtcttt	3480
ttttattgag	tctctctctt	tgctgatccc	ttccaagggg	tcaatatata	tacatgtata	3540
tactgtacat	atgtatatgt	aactaatata	catacatata	ggatatgtata	tgtaatgggt	3600
atatgtactc	atgttccttg	tgtagcaacg	tgtgggtatg	ctacacagag	aacatgagaa	3660
cataaagcca	tttttatgct	tactactaaa	agctgtccac	tgtagagttg	ctgtatgtag	3720
caatgtgtat	ccactctaca	gtggctcagct	tttagtagag	agcataaaaa	tgataaaaata	3780
cttcttga aa	acttagttta	ctatacatct	tgccctatta	atatgttctc	ttaacgtgtg	3840
ccattgttct	ctttgacctat	tttctatata	tgatgttgat	gttcaaacacc	tggactgaat	3900
gtctgttctc	agatcccttg	gatgttacag	atgaggcagt	ctgactgtcc	tttctacttg	3960
aaagattaga	atatgtatcc	aaatggcatt	cacgtgtcac	ttagcaagggt	ttgctgatgc	4020
ttcaaagagc	ttagtttgcg	gtttcctgga	cgtggaaaca	agtatctgag	ttccctggag	4080
atcaacggga	tgaggtgtta	cagctgcctc	cctcttcatg	caatctgggtg	agcagtggtg	4140
caggcgggga	gccagagaaa	cttgccagtt	atataacttc	tctttggctt	ttcttcatct	4200
gtaaaacaag	gataatactg	aactgtaagg	gttagtggag	agtttttaat	taaaagaatg	4260
tgtgaaaagt	acatgacaca	gtagttgctt	gataatagtt	actagtagta	gtattcttac	4320
taagacccaa	tacaaatgga	ttattttaac	caaaaaaaaa	aaaaaaaaaa	aa	4372

<210> 240

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 240

agttcagaca ctgtactcga agtgggaagga ctttcatttt gagaagatac catttgatcc 60

<210> 241

<211> 1647

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 241

```

gcgcgcggggt ttcgttgacc cgcggcgttc acgggaattg ttcgctttag tgccggcgcc 60
atgggggtcgg agctgactcg gcgcctagcc ccgcgcctgg gcctcgccga gcccgacatg 120
ctgaggaaag cagaggagta cttgcgcctg tcccgggtga agtgtgtcgg cctctccgca 180
cgcaccacgg agaccagcag tgcagtcatg tgcctggacc ttgcagcttc ctggatgaag 240
tgccccttgg acagggtcta ttttaattaaa ctttctgggt tgaacaagga gacatatcag 300
agctgtctta aatctt ttga gtgtttactg ggctgaatt caaatattgg aataagagac 360
ctagctgtac agtttagctg tatagaagca gtgaacatgg cttcaaagat actaaaaagc 420
tatgagtcca gtcttccca gacacagcaa gtggatcttg acttatccag gccactt ttc 480
acttctgctg cactgtcttc agcatgcaag attctaaagc tgaaagtgga taaaaacaaa 540
atggtagcca catccggtgt aaaaaaagct atatttgatc gactgtgtaa acaactagag 600
aagattggac agcaggctcg cagagaacct ggagatgtag ctactccacc acggaagaga 660
aagaagatag tggttgaagc ccagcaaaag gaaatggaga aggtagagga gatgcca cat 720
aaaccacaga aagatgaaga tctgacacag gattatgaag aatggaaaag aaaaatt ttg 780
gaaaatgctg ccagtgtcta aaaggctaca gcagagtgat ttcagcttcc aaactggtat 840
acattccaaa ctgatagtac attgccatct ccaggaagac ttgacggctt tgggatt ttg 900
tttaaaacttt tataataagg atcctaagac tgttgctttt aaatagcaaa gcagcctacc 960
tggaggctaa gtctgggcag tgggctggcc cctgggtgtga gcattagacc agccacagt 1020
cctgattggg atagccttat gtgctttcct acaaaatgga attggaggcc gggcgcatg 1080
gctcacgcct gtaatccag cactttggga ggccaagggt ggtggatcac ctgaggtcag 1140
gagctcgaga ccagcctggc caacatgggt aaaccccatc tctactaaaa atacaaaat 1200
tagccagggt tgatggtgca tgcctgtaat ccagctcct cagtaggctg agacaggagc 1260
atcacttgaa cgtgggaggc agaggttgca gtgagccgag attgcaccac cgcactccag 1320
cctgggtgac agagcgagac ttatctcata aataaataga tagatactcc agcctgggtg 1380
acagagcgag acttatagat agatagatag atagatggat agatagatag atagatagat 1440
agatagataa acggaa ttgg agccattttg ctttaagtga atggcagtc cttgtcttat 1500
tcagaatata aaattcagtc tgaatggcat cttacagatt ttacttcaat ttttgtgtac 1560
ggtatttttt atttga ctaa atcaatatat tgtacagcct aagttaataa atgttat tta 1620
tatatgcaaa aaaaaa aaaa aaaaaa 1647

```

<210> 242

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 242

```

tgctttaagt gaatggcagt cccttgtctt attcagaata taaaattcag tctgaatggc 60

```

<210> 243

<211> 1455

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014364

<400> 243

```

ggcggtccgc acgcacctcg gtaacatcac agcagggtcca ggccaatgat aaccttataa 60
gaggccatgt cgaagcgcga catcgtcctc accaatgtca ccgttgtcca gttgctgcga 120
cagccgtgcc cggtagcag agcaccgccc ccacctgagc ctaaggctga agtagagccc 180
cagccacaac cagagccac accagtcagg gaggaaataa agccaccacc gccaccactg 240
cctcctcacc ccgtactcc tctcctaag atgggtgtctg tggcccgga gctgactgtg 300
ggcatcaatg gatttgagc catcggtcgc ctggctcctg gcgcctgcat ggagaagggt 360
gttaagggtg tggctgtgaa tgatccatc attgacccgg aatacatggt gtacatgttt 420
aagtatgact ccaccacgg ccgatacaag ggaagtgtgg aattcaggaa tggacaaactg 480
gtcgtggaca accatgagat ctctgtctac cagtgcaaag agcccaaaca gatccccctg 540

```

```

agggtctgtcg ggagccc cta cgtgggtggag tccacaggcg tgtacctctc catacaggca 600
gcttcgggacc acatctc tgc aggtgctcaa cgtgtgggtca tctccgcgcc ctcacaggat 660
gcaccaatgt tcgtcatggg tgtcaatgaa aatgactata accctgggtc catgaacatt 720
gtgagcaacg cgtcctgcac caccaactgt ttgggtcccc tcgccaaagt catccacgag 780
cgatttggga tcgtggaagg gttgatgacc acagtccatt cctacacggc caccagaag 840
acagtggacg ggccatcaag gaaggcctgg cgagatgggc ggggtgcca ccagaacatc 900
atcccagcct ccactggggc tgcgaaagct gtgaccaaag tcatcccaga gtcacaagg 960
aagctgacag ggatggcgtt ccgggtacca acccggatg tgtctgtcgt ggacctgacc 1020
tgccgcctcg cccagcc tgc cccctactca gccatcaagg aggtgtgtaa agcagcagcc 1080
aaggggcca tggctggcat ccttgccctac accgaggatg aggtcgtctc tacggacttc 1140
ctcggtgata cccactcgtc catcttcgat gctaaggccg gcattgcgct caatgacaat 1200
ttcgtgaagc tcatttcctg gtacgacaac gaatatggct acagtcaccg ggtggtcgac 1260
ctcctccgct acatgttcag ccgagacaag tgaaacggga aggtcctttc tttccttccc 1320
aggggcccgg gccggacat gtgcctcccg ttccagcatc tggtgcccg ggggagggaag 1380
gacacccggg gcgggcgccc cagccgatg ggtccatggt gaaataaaaa acagtgtctc 1440
aaaaaaaaa aaaaa 1455

```

```

<210> 244
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014364

```

```

<400> 244
cgctcaatga caatttcgtg aagctcattt catggtacga caacgaatat ggctacagtc 60

```

```

<210> 245
<211> 935
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014462

```

```

<400> 245
gaagtgggta agggtaatat ggaggagctt ccggcaggcc ccggcggctg aaagcggggg 60
cagaagtgtt ggtctcggtc gggattccgg gcttgggtccc accgaggcgg cgactgcggg 120
aggagggaag aggttttggg cgcgctggcc tcccgcgct gtgcattgca gcattatttc 180
agttcaaaat gaactatatg cctggcaccg ccagcctcat cgaggacatt gacaaaagc 240
acttggttct gcttcgagat ggaaggacac ttataggctt tttaagaagc attgatcaat 300
ttgcaaaact agtgctacat cagactgtgg agcgtattca tgtgggcaaa aaatacgggtg 360
atattcctcg agggattttt gtggtcagag gagaaaatgt ggtcctacta ggagaaatag 420
acttggaaaa ggagagtgc acaccctcc agcaagtatc cattgaagaa attctagaag 480
aacaaagggg ggaacagcag accaagctgg aagcagagaa gttgaaagtg caggccctga 540
aggaccgagg tctttccatt cctcgagcag atactcttga tgagtactaa tcttttgccc 600
agaggctgtt ggctcttgaa gtagtagggc tgtcactgag tgaaagtgc atcctggcca 660
cctcacgcac ttgatcacag actgtagagt ttgaaaagt cacttttatt tttaattatt 720
ttacatatgc aacatgaaga aatcgtgtag gtgggttttt ttttaataa caaaatcact 780
gtttaaagaa acagtggcat agactccttc acacatcact gtggcaccag caactacttc 840
tttatattgt tcttcataat ccaaattaga gtttacaggg acagtcttca tttacttgta 900
aataaaatat gaatctcaaa aaaaaaaaaa aaaaa 935

```

```

<210> 246
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014462

```

<400> 246
ttaaataacaa aatcactgtt taaagaaaca gtggcataga ctccttcaca catcactgtg 60

<210> 247
<211> 890
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014501

<400> 247
ggcggaccga agaacgcagg aagggggccg gggggacccg cccccggccg gccgcagcca 60
tgaactccaa cgtggagaac ctacccccc acatcatccg cctggtgtac aaggaggtga 120
cgacactgac cgcagaccca cccgatggca tcaagggtctt tcccaacgag gaggacctca 180
ccgacctcca ggtcaccatc gagggccctg aggggacccc atatgctgga ggtctgttcc 240
gcatgaaact cctgctgggg aaggacttcc ctgcctcccc acccaagggc tacttctctga 300
ccaagatctt ccacccgaac gtgggcgcca atggcgagat ctgcgtcaac gtgctcaaga 360
gggactggac ggcctgagctg ggcattccgac acgtactgct gaccatcaag tgcctgctga 420
tccaccctaa ccccgagtct gcactcaacg aggaggcggg ccgcctgctc ttggagaact 480
acgaggagta tgcggctcgg gcccgctctgc tcacagagat ccacgggggg gccggcgggg 540
ccagcggcag ggcgaagcc ggtcggggcc tggccagtgg cactgaagct tccctccaccg 600
accctggggc ccagggggg ccgggagggg ctgagggtcc catggccaag aagcatgctg 660
gcgagcgcga taagaagctg gcggccaaga aaaagacgga caagaagcgg gcgctgcggg 720
cgctgcggcg gctgtagtgg gctctcttcc tccttccacc gtgaccccaa cctctctctg 780
cccctccctc caactctgtc tctaagttat ttaaattatg gctgggggtcg gggagggtac 840
agggggcact gggacctgga tttgttttcc taaataaagt tggaaaagca 890

<210> 248
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014501

<400> 248
acacgtactg ctgacctca agtgctgtg gatccaccct aaccccgagt ctgcactcaa 60

<210> 249
<211> 1182
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016095

<400> 249
gcgccggcg gcgtctctc cgggacgct gaggggcccg aggagaccgt gaggctcttg 60
cctgcagctc gcgcggccat ggacgctgcc gaggtcgaat tcctcgccga gaaggagctg 120
gttaccatta tccccaactt cagtctggac aagatctacc tcctcggggg ggaacctggg 180
ccttttaacc ctgggtttacc cgtggaagtg cccctgtggc tggcgattaa cctgaaacaa 240
agacagaaat gtcgctgct ccctccagag tggatggatg tagaaaagtt ggagaagatg 300
agggatcatg aaCGaaagga agaaactttt accccaatgc ccagccctta ctacatggaa 360
cttacgaagc tccgtgttaa tcattgcttca gacaacatcc cgaaggcaga cgaaatccgg 420
accctggtca aggatatgtg ggacactcgt atagccaaac tccgagtgtc tgcctgacagc 480
tttgtgagac agcaggaggc acatgccaaag ctggataact tgaccttgat ggagatcaac 540
accagcggga ctttctctcac acaagcgctc aaccacatgt acaaactccg cacgaacctc 600

```

cagcctctgg agagtactca gtctcaggac ttctagagaa aggcctgggtg caggcgggctt 660
gctgggggat gtgagcgctc aggatgtgat gaggtactcg tggttcttga gctctagaaa 720
cacttctgat gcatgaaaaa tgtgtgatgg tgcaagggaat ggattcagga tgttgttga 780
gaaacaagtt tgtgattagt ccttaaaact tagctccctg ggacattctt caattccaca 840
tctgtttcta gaaaccagcc ctttttcccc ccacttttga gaaataaaaa agccttaggt 900
aaataagtca ttctccctag cagagccact tgggtctcct gcatggaagc cgtcacactt 960
gggcaggtgt tcagtactg gttaggtgtag atacagcagg agtggccatg tgggtccacgg 1020
ctttttaccc cttcttgatc ctgatttctt gggctgaatt tagactctct cacagaggtg 1080
gctcacagag aaggatggca gatggtgcag ccaacaatgc tgaccggtgc ttaacctcta 1140
agccctgatc cacaataaaa atggacccaa ctcaaaaaaa aa 1182

```

<210> 250

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 250

```

atggattcag gatgttgttg gagaaacaag tttgtgatta gtacctaaaa cttagctccc 60

```

<210> 251

<211> 704

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016185

<400> 251

```

tgcagcgggtg gtcggtctgtt ggggtgtggag tttccagcgg cccctcgggt ccgacccttt 60
gagcgttctg ctccggcgcc agcctacctc gctcctcggc gccatgacca caaccaccac 120
cttcaaggga gtcgacccca acagcaggaa tagctccga gttttgcggc ctccagggtg 180
tggatccaat ttttcattag gttttgatga accaacagaa caacctgtga ggaagaacaa 240
aatggcctct aatatctttg ggacacctga agaaaatcaa gcttcttggg ccaagtcagc 300
aggtgccaaag tctagtgggtg gcaggggaaga cttggagtca tctggactgc agagraaggaa 360
ctcctctgaa gcaagctccg gagacttctt agatctgaag ggagaagggtg atat tcatga 420
aaatgtggac acagacttgc caggcagcct ggggcagagt gaagagaagc ccgtgcctgc 480
tgogcctgtg cccagccccgg tggccccggc cccagtgcc tccagaagaa atccacctgg 540
cggcaagtcc agcctcgtct tgggttagct ctgactgtcc tgaacgctgt cgttctgtct 600
gtttcctoca tgcttgagaa ctgcacaact tgagcctgac tgtacatctt cttggatttg 660
tttcattaaa aagaagcact ttatgtaaaa aaaaaaaaaa aaaa 704

```

<210> 252

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016185

<400> 252

```

tgaaccaaca gaacaacctg tgaggaagaa caaaatggcc tctaatatct ttgggacacc 60

```

<210> 253

<211> 2268

<212> DNA

<213> Homo sapiens

<220>

<221> Modified_base
 <222> 1 ... 2268
 <223> n = a,c,g, or t

<300>
 <308> NM_016359

<400> 253
 gggatttgaa ccncgctgac gaagtttggt gatccatctt ccgagtatcg ccgggatttc 60
 gaatcgcgat gatcatcccc tctctagagg agctggactc cctcaagtac agtgacctgc 120
 agaacttagc caagagtctg ggtctccggg ccaacctgag ggcaaccaag ttgttaaaag 180
 ccttgaaagg ctacattaaa catgaggcaa gaaaaggaaa tgagaatcag gatgaaagtc 240
 aaacttctgc atcctcttgt gatgagactg agatacagat cagcaaccag gaagaagctg 300
 agagacagcc acttggccat gtcacaaaa caaggagaag gtgcaagact gtccgtgtgg 360
 accctgactc acagcagaat cattcagaga taaaaataag taatccact gaattccaga 420
 atcatgaaaa gcaggaaagc caggatctca gagctactgc aaaagttcct tctccaccag 480
 acgagcacca agaagctgag aatgctgttt cctcaggtaa cagagattca aaggtacctt 540
 cagaaggaaa gaaatctctc tacacagatg agtcatccaa acctggaaaa aataaaagaa 600
 ctgcaatcac tactc caaac tttaagaagc ttcatgaagc tcattttaag gaaatggagt 660
 ccattgatca atata ttgag agaaaaagaa acattttgaa gaacacaatt ccatgaatga 720
 actgaagcag cagcccatca ataaggagg ggtcaggact ccagtacctc caagaggaag 780
 actctctgtg gcttctactc ccatcagcca acgacgctcg caaggccggg cttgtggccc 840
 tgcaagtcag agtaccttgg gtctgaagggt gtcactcaag cgctctgcta tctctgcagc 900
 taaaacgggt gtcaggtttt cagctgtctac taaagataat gagcataagc gttcactgac 960
 caagactcca gccagaaagt ctgcacatgt gaccgtgtct gggggcacc caaaaggcga 1020
 ggctgtgctt gggacacaca aattaaagc catcacgggg aattctgctg ctgttattac 1080
 ccatttcaag ttgacaactg aggcaacgca gactccagtc tccaataaga aaccagtgtt 1140
 tgatcttaaa gcaagtttgt ctctgtccct caactatgaa ccacacaaag gaaagctaaa 1200
 accatggggg caatctaaag aaaataatta tctaaatcaa catgtcaaca gaattaaactt 1260
 ctacaagaaa acttacaac aaccccatct ccagacaaag gaagagcaac ggaagaaacg 1320
 cgagcaagaa cgaaaggaga agaaagcaaa ggttttgagg atgcgaaggg gcctcatttt 1380
 ggctgaagat taataatttt ttaatatctt gtaaataattc ctgtattctc aacttttttc 1440
 cttttgtaaa tttttttttt tttgctgtca tccccacttt agtcacgaga tctttttctg 1500
 ctaactgttc atagtctgtg tagtgtccat ggggttcttca tgtgctatga tctctgaaaa 1560
 gacgttatca ccttaaagct caaattcttt gggatgggtt ttacttaagt cattaacaa 1620
 ttcaggtttc taacgagacc catcctaaaa ttctgtttct agatttttaa tgtcaagttc 1680
 ccaagttccc cctgctgggt ctaatatata cagaactgca gtcttctgct agccaatagc 1740
 atttacctga tggcagctag ttatgcaagc ttcaggagaa tttgaacaat aacaagaata 1800
 gggtaagctg ggatagaaag gccacctctt cactctctat agaatatagt aacctttatg 1860
 aaacggggcc atatagtttt gttatgacat caatatttta cctaggtgaa attgtttagg 1920
 cttatgtacc ttctgtcaaa tatcctcatg taattgccat ctgtcactca ctatattcac 1980
 aaaaataaaa ctctacaact cattctaaca ttgcttactt aaaagctaca tagccctatc 2040
 gaaatgcgag gattaatgct ttaatgcttt tagagacagg gtctcactgt gttgcccagg 2100
 ctggtctcaa actccaccaa atgtacttct tattcatttt atggaaaaga ctaggctttg 2160
 cttagtatca tgtccatggt tccttcacct cagtggagct tctgagtttt atactgctca 2220
 agatcgtcat aaataaaatt ttttctcatt gtcaaaaaaa aaaaaaaa 2268

<210> 254
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_016359

<400> 254
 acattgctta cttaaaagct acatagccct atcgaaatgc gaggattaat gctttaatgc 60

<210> 255
 <211> 1590
 <212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 2 55

```

gaggcagttc tgttgccact ctctctcctg tcaatgatgg atctcagaaa taccaccagcc 60
aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcgaccatg ccattgacat catctgtggg ttctgaagg aaaggtgctt ccgaggtagc 180
tctaccctg tgtgtgtgtc caaggtggta aaggtggct cctcaggcaa gggcaccacc 240
ctcagaggcc gatctgacgc tgacctgggt gtcttcctca gtcctctcac cacttttcag 300
gatcagttaa atcgccgggg agagttcatc caggaaatta ggagacagct ggaagcctgt 360
caaagagaga gagcactttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420
cgtgcgctca gcttcgtact gagttcgctc cagctcgggg agggggtgga gttcgatgtg 480
ctgctgcct ttgatgccct gggtcagttg actggcagct ataaacctaa ccccaaatc 540
tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600
ttcacagaac tacagagaga cttcctgaag cagcgcccca ccaagctcaa ggcctcatc 660
cgctagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720
tatgcccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780
aacacagccc aaggatttctg gacggtcttg gaattagtca taaactacca gcaactctgc 840
atctactgga caaagtatta tgactttaaa aacccccatta ttgaaaagta cctgagaagg 900
cagctcacga aaccagggcc tgtgatcctg gaccgcggcg accctacagg aaacttgggt 960
ggtggagacc caaagggttg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020
ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggctgaaagc 1080
aacagtaacg acgatgagac cgacgatccc aggcagctatc agaaatatgg ttacattgga 1140
acacatgagt accctcattt ctctcataga ccagcacgc tccaggcagc atccaccca 1200
caggcagaag aggactggac ctgcaccatc ctctgaatgc cagtgcattt tgggggaaag 1260
ggctccagtg ttatctggac cagttccttc attttcaggt gggactcttg atccagagaa 1320
gacaaagctc ctcaagtgagc tgggtgtataa tccaagacag aaccctaagtc tctgactcc 1380
tggccttctc tgccctctat cctatcatag ataacattct ccacagcctc acttcattcc 1440
acctattctc tgaaaatatt ccctgagaga gaacagagag atttagataa gagaatgaaa 1500
ttccagcctt gactttcttc tgtgcacctg atgggagggt aatgtctaata gtattatcaa 1560
taacaataaa aataaagcaa ataccaaaaa 1590

```

<210> 25 6

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 25 6

```

cgatcccaagg acgtatcaga aatatgggta cattggaaca catgagtacc ctcatctctc 60

```

<210> 25 7

<211> 29 05

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 25 7

```

cggcagccag ctgagagcaa tgggaaatgg ggagtcccag ctgtcctcgg tgctgtctca 60
gaagctgggt tggtttatcc aggaatacct gaagccctac gaagaatgtc agacactgat 120
cgacgagatg gtgaacacca tctgtgacgt ctgcaggaac ccgaacagt tccccctggt 180
gcaggagtg gccatagggt gctcctatgg acggaaaaca gtcttaagag gcaactccga 240
tggtagccctt gtccttttct tcagtgaactt aaaacaattc caggatcaga agagaagcca 300
acgtgacatc ctcgataaaa ctggggataa gctgaagttc tgtctgttca cgaagtgggt 360

```

gaaaaacaat	tctcgagatcc	agaagtcctt	tgatgggtcc	accatccagg	tggtcacaaa	420
aatcagaga	atctcttttcg	aggtgctggc	cgccttcaac	gctctgagct	taaatgataa	480
tcccagcccc	tggatctatc	gagagctcaa	aagatccttg	gataagacaa	atgccagtcc	540
tggtgagttt	gcagtcctgct	tcactgaact	ccagcagaag	ttttttgaca	accgtcctgg	600
aaaactaaag	gatttgatcc	tcttgataaa	gcactggcat	caacagtgc	agaaaaaat	660
caaggattta	ccctcgctgt	ctccgtatgc	cctggagctg	cttacgggtg	atgcctggga	720
acaggggtgc	agaaaagaca	actttgacat	tgctgaaggc	gtcagaacgg	ttctggagct	780
gatcaaatgc	caggagaagc	tgtgtatcta	ttggatgggc	aactacaact	ttgaagatga	840
gaccatcagg	aacatcctgc	tgcaccagct	ccaatcagcg	aggccagtaa	tcttggatcc	900
agttgaccca	accaataatg	tgagtggaga	taaaatatgc	tggcaatggc	tgaaaaaga	960
agctcaaacc	tggttgactt	ctcccaacct	ggataatgag	ttacctgcac	catcttggaa	1020
tgtcctgcct	gcaccactct	tcacgacccc	aggccacctt	ctggataagt	tcatcaagga	1080
gtttctccag	ccaacaaat	gcttctctaga	gcagattgac	agtgcctgta	acatcatccg	1140

tacattcctt	aaagaaaact	gcttccgaca	atcaacagcc	aagatccaga	ttgtccgggg	1200
aggatcaacc	gccaaaggca	cagctctgaa	gactggctct	gatgccgatc	tcgtcgtgtt	1260
ccataactca	cttaaaagct	acacctccca	aaaaaacgag	cggcacaaaa	tcgtcaagga	1320
aatccatgaa	cagctgaaag	ccttttggag	ggagaaggag	gaggagcttg	aagtgcagctt	1380
tgagcctccc	aagtgggaag	ctcccagggt	gctgagcttc	tctctgaaat	ccaaagtcct	1440
caacgaaagt	gtcagctttg	atgtgcttcc	tgcctttaat	gcactgggtc	agctgagttc	1500
tggctccaca	ccagccccg	aggtttatgc	agggctcatt	gatctgtata	aatcctcgga	1560
cctcccggga	ggagagtttt	ctacctgttt	cacagtcctg	cagcgaaaact	tcattcgctc	1620
ccggcccacc	aaactaaagg	atttaattcg	cctgggtgaag	cactgggtaca	aagagtgtga	1680
aaggaaaactg	aagccaaagg	ggtcctttgcc	cccaaagtat	gccttggagc	tgctcaccat	1740
ctatgcctgg	gagcagggga	gtggagtgcc	ggattttgac	actgcagaag	gtttccggac	1800
agtcctggag	ctgggtcacac	aatatcagca	gctcggcatc	ttctgggaag	tcaattacaa	1860
ctttgaagat	gagaccgtga	ggaagtttct	actgagccag	ttgcagaaaa	ccaggcctgt	1920
gatcttggac	ccaggcgaac	ccacaggtga	cgtgggtgga	ggggaccggt	ggtgttggca	1980
tcttctggac	aaagaagcaa	aggttagggt	atcctctccc	tgcttcaagg	atgggactgg	2040
aaacccaata	ccaccttggg	aagtgccgac	aatgcagaca	ccaggaagtt	gtggagctag	2100
gatccatcct	attgtcaatg	agatgttctc	atccagaagc	catagaatcc	tgaataataa	2160
ttctaaaaga	aacttctgga	gatcatctgg	caatcgcttt	taaagaactcg	gtccaccgtg	2220
agaaagagtc	actcacatcc	attcttccct	tgatggctcc	tattctctct	tccttgcct	2280
tcttggactt	cttgaaatca	atcaagactg	caaacccttt	cataaagctg	ccttgcctgaa	2340
ctcctctctg	caggagccct	gcttaaaata	gttgatgtca	tcactttatg	tgcatcttat	2400
ttctgtcaac	ttgtattttt	ttttcttgta	tttttccaat	tagctcctcc	tttttccctc	2460
cagtctaaaa	aaggaatcct	ctgtgtcttc	aaagcaaagc	tctttacttt	cccttgggtt	2520
ctcataactc	tgtgatcttg	ctctcggtgc	ttccaactca	tccacgtcct	gtctgtttcc	2580
tctgtatata	aaaccttttc	tgccccgtct	gacacagaca	tcctctatgc	cagcagccag	2640
gccaaacctt	tcattagaac	ttcaagctct	ccaaaggctc	agattataac	tgttgcata	2700
tttatatgag	gctgttgcct	tttcttctg	agcctgcctt	tatcccccca	ccaggagta	2760
tcctcttgcc	aaagcaaaag	actttttcct	tggctttagc	cttaaagata	cttgaaggtc	2820
taggtgcttt	aacctcacat	acctcactt	aaacttttat	cactgttgca	tataccagtt	2880
gtgatacaat	aaagaatgta	tctgg	2905			

<210> 258

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 258

aaggtctagg tgctttaacc tcacataccc tcacttaaac ttttatcact gttgcatata 60

<210> 259

<211> 2054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 259

```

gggaagctcg ggcgggcagg gtttccccgc acgctggcgc ccagctcccg gcgcggaggc 60
cgctgtaagt ttcgctttcc attcagtggg aaacgaaagc tgggcggggt gccacgagcg 120
cggggccaga ccaaggcggg cccggagcgg aacttcgggt ccagctcggg ccccggtca 180
gtcccagcgt ggaactcagc agcggaggct ggacgcttgc atggcgcttg agagattcca 240
tcgtgcctgg ctcacataag cgcttcctgg aagtgaagtc gtgctgtcct gaacgcgggc 300
caggcagctg cggcctgggg gttttggagt gatcacgaat gagcaaggcg tttgggctcc 360
tgaggcaaatt ctgtcagtc atcctggctg agtccctgca gtccccggca gatcctgaag 420
aaaagaagga agaagacagc aacatgaaga gagagcagcc cagagagcgt cccagggcct 480
gggactaccc tcatggcctg gttggtttac acaacattgg acagacctgc tgccttaact 540
ccttgattca ggtgttcgta atgaatgtgg acttcaccag gatattgaag aggatcacgg 600
tgcccagggg agctgacgag cagaggagaa gcgtcccttt ccagatgcct ctgctgctgg 660
agaagatgca ggacagccgg cagaaagcag tgcggccctt ggagctggcc tactgcctgc 720
agaagtgcaa cgtgcccttg tttgtccaac atgatgctgc ccaactgtac ctcaaactct 780
ggaacctgat taaggaccag atcactgatg tgcacttggg ggagagactg caggccctgt 840
atacgatecg ggtgaaggac tccttgattt gcgttgactg tgcctatggg agtagcagaa 900
acagcagcat gctcaccctc ccactttctc tttttgatgt ggactcaaag cccctgaaga 960
cactggagga cgccttcac tgcttcttcc agcccaggga gttatcaaagc aaaagcaagt 1020
gcttctgtga gaactgtggg aagaagaccc gtgggaaaca ggtcttgaag ctgaccatt 1080
tgcccagac cctgacaatc cacctcatgc gattctccat caggaattca cagacgagaa 1140
agatctgcca tccctgtac tcccccaga gcttggattt cagccagatc cttccaatga 1200
agcgagagtc ttgtgatgtc gaggagcagt ctggaggggca gtatgagctt tttgctgtga 1260
ttgcgcacgt gggaatggca gactccgggc attactgtgt ctacatccgg aatgctgtgg 1320
atggaaaatg gttctgcttc aatgactcca atatttgctt ggtgtcctgg gaagacatcc 1380
agtgtacctc cggaaatcct aactaccact ggcaggaaac tgcatacttt ctggtttaca 1440
tgaagatgga gtgctaattg aaatgcccaa aaccttcaga gattgacaag ctgtcatttt 1500
ccatttcctg tccctggatc acggagtctt ctaagagatt ttgcaatgag gagaagcatt 1560
gttttcaaac tatataactg agccttattt ataattaggg atattatcaa aatatgtaac 1620
catgaggccc ctacaggtcct gatcagtcag aatggatgct ttcaccagca gaccgggcca 1680
tgtggctgct cggtcctggg tgctcgctgc tgtgcaagac attagcccct tagttatgag 1740
cctgtgggaa cttcaggggt tccagtgagg gagagcagtg gcagtgggag gcactctgggg 1800
gccaaaggtc agtggcaggg ggtatttcag tattatacaa ctgctgtgac cagacttgta 1860
tactggctga atatcagtg tgtttgtaat ttttcacttt gagaaccaac attaattcca 1920
tatgaatcaa gtgtttttgta actgctattc atttattcag caaatattta ttgatcatct 1980
cttctccata agatagtgtg ataaacacag tcatgaataa agttattttc cacaaaaaaa 2040
aaaaaaaaaa aaaa 2054

```

<210> 260

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 260

```

tgagcatctc ttctccataa gatagtgtga taaacacggg catgaataaa gttattttcc 60

```

<210> 261

<211> 3638

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017523

<400> 261

```

ggtagatgcg gctgtgacag cagcaaagaa tgacggccaa gggcgacagc aggggctggc 60

```

catgctgtaa	aggggcttct	tgggaggggtc	cagcctcagg	aatcaagggg	aactcctgag	120
ccgagaattc	tgaagatctc	ctccctccct	gaagctgtgg	gctgggcat	cggaaaactt	180
tcagttttgt	ttccttgect	gcaagaaacg	aaactcaacc	gaaagcctgc	agagagcaga	240
acatggaagg	agactttctg	gtgtgcagga	actgtaaaag	acatgtagtc	tctgccaaact	300
tcaccctcca	tgaggcttac	tgcctgcggt	tcctggctct	gtgtccggag	tgtgaggagc	360
ctgtcccca	ggaaaccatg	gaggagcaact	gcaagcttga	gcaccagcag	gttggtgtga	420
cgatgtgtca	gcagagcatg	cagaagtcct	cgctggagtt	tcataaggcc	aatgagtgcc	480
aggagcgccc	tgttgagtgt	aagttctgca	aactggacat	gcagctcagc	aagctggagc	540
tccacgagtc	ctactgtggc	agccggacag	agctctgcc	aggctgtggc	cagttcatca	600
tgcaccgcat	gctcgccag	cacagagatg	tctgtcgcag	tgaacaggcc	cagctcgga	660
aaggggaaag	aatttcagct	cctgaaagg	aaatctactg	tcattattgc	aaccaaata	720
ttccagaaaa	taagtatttc	caccatatgg	gtaaatgttg	tccagactca	gagtttaaga	780
aacacttttc	tgttggaat	ccagaaattc	ttccttcac	tcttccaagt	caagctgctg	840
aaaatcaaac	ttccacgatg	gagaaagatg	ttcgtccaaa	gacaagaagt	ataaacagat	900
ttcctcttca	ttctgaaagt	tcatacaaga	aagcaaaaac	aaaaccttgg		960
atccactttt	gatgtcagag	cccaagccca	ggaccagctc	ccctagagga	gataaagcag	1020
cctatgacat	tctgaggaga	tgttctcagt	gtggcatcct	gcttccctg	ccgatccctaa	1080
atcaacatca	ggagaaatgc	cggtgggttag	cttcatcaaa	aggaaaaaa	gtgagaaatt	1140
tcagctagat	ttggaaaagg	aaaggacta	caaattcaaa	agatttcaact	tttaacactg	1200
gcattcctgc	ctacttgctg	tgggtggtctt	gtgaaagggtg	atgggtttta	ttcgttgggc	1260
tttaaaagaa	aagggttggc	agaactaaaa	acaaaactca	cgtatcatct	caatagatac	1320
agaaaaggct	tttgataaaa	ttcaacttga	cttcatgtta	aaaacctca	acaaaccagg	1380
cgctgaagga	acatacctca	aaataataag	agccatctat	gacaaaac	caagccaacat	1440
catactgaat	gagcaaaaagc	tggagcatta	ctcttgagaa	gtagaacaag	gcacttcagt	1500
cctattcaac	atagtactgg	aagtctctgc	cacagcaatc	aggcaaga	gaaagataaa	1560
aggcaaccaa	aaagaaagga	agtcgaagta	tctctgtttg	cagacgat	gattctatat	1620
ctagaaaacc	c catgatctt	ggcccaaaag	ctcctagatc	tgataaac	cttcagctaa	1680
ctttcaggag	acaaaatcaa	tatacaaaat	atggtagcat	ttttatacac	caacgacatc	1740
caagctgaga	gccaaatcaa	gaatgcaatc	ctattcacaa	ttgccacaaa	aagaataaaa	1800
tacctaggaa	tacagctaac	cagggagatg	aaagatctct	acaacaaa	ttacaaaaca	1860
ctgctgaaag	aaatcagaga	tgacacaaat	ggaaaaacat	tccatact	tgataggaa	1920
gaatcaatat	tgtttaaagt	gccatactac	ccaagcaat	ttatagat	aatgctattc	1980
ctatcaaact	accaataaca	ttcttcacag	aatcagaaaa	aaaaagca	aaaatttatt	2040
tgaaacccaa	aaagagccca	aaaagcccaa	gcaatcctaa	gcaaaaagaa	caaagctgga	2100
ggcatcgcat	tacccaactt	caaactatac	tacagggcta	cagtaacc	aactgcatga	2160
tactggtaca	aaagcatggt	gctggtacaa	aagcagacac	atagatcaat	ggaacagaat	2220
agagggccca	gaaataaagc	tacacaccta	caaccatcta	atctttgaca	aagttgacaa	2280
aaatacgcaa	tggggaaaga	attccccatt	cagtaagtgg	tactggga	actagctagc	2340
catatgcaga	ggattgaaac	tgaaccactt	ccttacacca	tatgcaaaa	tcaactcaag	2400
atggattaaa	gacttaaatg	taaaacccca	aactataaaa	actctggaag	ataacctagg	2460
caataccatt	ctgggacatag	gaacggaaaa	agatttcatg	acaaagatcc	caaaaataat	2520
tgtaacgaaa	gcaaaaattg	acaaatggga	catgattaaa	cagaattacc	atttgactca	2580
gcaatcccat	tattgggttat	atacccaaag	gaatctaaat	cattctgtca	taaagacata	2640
tatacacaaa	tgttcacggc	agcactatac	acaatcgcaa	agtcagggaa	tcaaaactaaa	2700
tgtccatcag	tggtagaaag	gataaagaaa	atgtggtggc	agggagtgg	ggctcatgtc	2760
tgtaatccca	gcactttggg	aggctgaggg	gggtggttca	cctgaggtca	ggagtttgag	2820
accagcctgg	cCaacatggc	gaaactcctg	ctccgctaaa	aatacgaaaa	ttagccaggc	2880
gtggtggcga	gcacctgtca	tcccagctac	ttgggaggcc	taggcgtgag	aatcgcttga	2940
acctggaagg	tgggtggttc	agtgagccga	gatcctgcca	ctgcactcca	gcctgggcaa	3000
ccaagcgaga	ctctgcctta	aaaaaaaaaa	aaagaaaatg	tggcacatat	acaccatgga	3060
atactatgca	gcataaaaaa	agaatgggat	catgtcctgt	gcagcaacgt	ggatggagct	3120
ggaagccatt	atcctaaatg	aactcactca	gaaacagaaa	accaaatacc	acatgttctc	3180
acttataagt	agaagctaaa	cattgagtac	acatggatac	aaagaaggga	accgcagaca	3240
ctggggccta	cctgaggtcg	gagcatggaa	ggaggggtgag	gatcaaaaaa	ctacctatct	3300
ggtactatgc	tttttatctg	gatgatgaaa	taatctgtac	aacaaacctt	ggtgacatgc	3360
aatttaccta	tatagcaagc	ctacacatgt	gcccctgaac	ctaaaaaaa	agttaaaaga	3420
aaaacgtttg	gattattttc	cctcttttga	acaaagacat	tggtttgc	aaggactaca	3480
ataaaaccaa	cgggaaaaaa	gaaaggttcc	agttttgtct	gaaaattctg	attaagcctc	3540
tgggcccctac	acctggagaa	tcctacaccc	acagaacccg	gctttgtccc		3600
caaagaataa	aaacacctct	ctaaaaaaa	aaaaaaa			3638

<210> 262
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_017523

<400> 262
 ttggaaaagg aaaggtacta caaattcaaa agatttcact ttttaacactg gcatttcctgc 60

<210> 263
 <211> 2461
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_018410

<400> 263
 atgctgggta cgctgcgcgc catggagggc gaggacgtgg aagacgacca gctgctgcag 60
 aagctcaggg ccagtcgccg ccgcttccag aggcgcgtgc agcggctgat agagaagtac 120
 aaccagccct tcgaggacac ccgggtgggtg caaatggcca cgctgacctc cgagacgcca 180
 cagggattga gaatttgggg tggaagacta ataaaggaaa gaaacaaagg agagatccag 240
 gactcctcca tgaagcccg gcacaggaca gatggctccg tgcaagctgc agcctgggggt 300
 cctgagcttc cctgcaccg cacagtcctg ggagccgatt caaaaagcgg tgaggctgat 360
 gccacgtcag accaggaaga gtcagttgct tgggccttag cacctgcagt gcctcaaagc 420
 cttttgaaaa atgaattaag aaggaaatac ttgacccaag tggatatact gctacaagg 480
 gcagagtatt ttgagtgtgc aggtaacaga gctggaaggg atgtacgtgt gactccgctg 540
 ccttcactgg cctcacctgc cgtgcctgcc ccgggatact gcagtcgtat ctccggaaag 600
 agtcctgggtg acccagcgaa accagcttca tctccagag aatgggattc tttgcatcct 660
 tcctccacag acatggcctt agtacctaga aatgacagcc tctccctaca agagaccagt 720
 agcagcagct tcttaagcag ccagcccttt gaagatgatg acatttgrca tgtgaccatc 780
 agtgacctgt acgcagggat gctgcactcc atgagccggc tgttgagcac aaagccatca 840
 agcatcatct ccaccaaac gttcatcatg caaaactgga actgcagrag gaggacatca 900
 tataagagca ggatgaacaa aacatattgc aaaggagcca gacgttc tca gaggagctcc 960
 aaggagaact tcataccctg ctctgagcct gtgaaaggga caggggcatt aagagattgc 1020
 aagaacgtat tagatgtttc ttgccgtaag acaggtttta aattgga aaa agcttttctt 1080
 gaagtcaaca gaccccaaat ccataagtta gatccaagtt ggaaggagcg caaagtgaca 1140
 ccctcgaagt attcttcctt gatttacttc gactccagtg caacata taa tcttgatgag 1200
 gaaaatagat ttaggacatt aaaatggtta atttctcctg taaaaat agt ttccagacca 1260
 acaatacgac agggccatgg agagaaccgt cagagggaga ttgaaat ccg atttgatcag 1320
 cttcatcggt aatatgtcct gagtcccagg aaccagcctc gccggat gtg cctccgggac 1380
 tcctggggcca tgaacatgta cagagggggg ctgctgagtc ctggtggcct tcagggctta 1440
 gaaaccgcga ggctgagttt acctccagc aaagcaaaag caaaaagt ttt aagtgaggct 1500
 tttgaaaacc taggcaaaaag atctctggaa gcaggtaggt gcctgcc caa gagcgattca 1560
 tcttcatcac ttcctaaagac caaccccaca cacagcgcaa ctcgccc gca gcagacatct 1620
 gaccttcacg ttcagggaaa tagttctgga atatttagaa agtcagt gtc acccagcaaa 1680
 actctttcag tcccagataa agaagtgcc ggccacggaa ggaatcgtta cgatgaaatt 1740
 aaagaagaat ttgacaagct tcatcaaaag tattgcctca aatctcc tgg gcagatgaca 1800
 gtgcctttat gtattggagt gtctacagat aaagcaagta tggaggt tgc atatcaaaa 1860
 gaaggcttct taggaaaatt aaatccagac cctcacttcc aggggtt cca gaagtgtcca 1920
 tcatcacccc tgggtgcag aaaaagtcta ctgggtcaa ctgcaat tga ggctccttca 1980
 tctacatgtg ttgctcgtgc catcacgagg gatggcacga gggacca tca gttccctgca 2040
 aaaagaccca ggctatcaga accccagggc tccggacgcc agggcaa ttc cctgggtgcc 2100
 tcagatgggg tggacaacac cgtcagaccg ggagaccagg gcagctc ttc acagcccaac 2160
 tcagaagaga gaggagagaa cacgtcttac aggatggaag agaaaagtga tttcatgcta 2220
 gaaaaatttg aaactaaaag tgtgtagcta ggttatttcg gagtgttatt tatcttcca 2280
 cttgctctct gtttgtattt ttgttttgtt tttgattctt gagactgtga ggacttggtt 2340
 gacttctctg cccttaaagt aaatattagt gaaattgggt ccatcagaga taacctcgag 2400
 ttcttggtgt agaaattatg tgaataaagt tgctcaatta gaaaaaaa aaaaaaaaaa 2460

a 2461

<210> 264

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018410

<400> 264

agtgatttca tgctagaaaa attggaaact aaaagtgtgt agctaggtta tttcggagtg 60

<210> 265

<211> 1405

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 265

cacctcgctc	gcagcctccc	cagcgcagca	gcccggctgt	gggcctgcgg	cagccgggtc	60
ttcctgggtc	ccacctcctg	gggcccacgg	gcggcaggaa	ggggctcggc	gggacgcgcc	120
gtcagggacc	tgaggaggaa	caacggaacg	cgttcggaac	ggcctggact	cccagactc	180
acccgactcg	tggccacacc	gggagaactg	aagcggcagt	agccggcgga	gacgcccagc	240
ccgaaggccg	gctgctaggg	agcagacagc	tgaaccgctt	gccagacgcc	gaaacccagt	300
gacgcccctc	accgctccac	cgtgctcccg	gctccccgcc	cccgccgcc	gcggggcccca	360
aggcgcgatg	gdcgcctgtc	ctggaggggc	ccatttccgt	ccgtcgtggg	gggaggcaca	420
gtgagtccac	tggggcacgg	cagcgtctaa	gccacaagcc	gagcacataa	gccaggtcct	480
aacggagcct	atgtgtaagt	ccactactgg	tgcaagggtt	cacacttcta	agaagagcgg	540
cgtggggggc	tcggcgacct	tcgcttcagt	cgctcccccg	tgcagtcccc	tgtgccaag	600
acacagcctg	atgcttgtgc	tccggtgggc	ggagcttgga	ggcggcggga	actgcaattg	660
gtggctttga	aggcgcggcg	agcgggaaca	gctcttgagg	agtgagactg	caggagatgt	720
gggcccgtgc	aaagagatgg	atgagactgt	tgctgagttc	atcaagagga	ccatcttgaa	780
aatccccatg	aatgaactga	caacaatcct	gaaggcctgg	gattttttgt	ctgaaaaatca	840
actgcagact	gtaaatttcc	gacagagaaa	ggaatctgta	gttcagcact	tgatccatct	900
gtgtgaggaa	aagcgtgcaa	gtatcagtga	tgctgcctcg	ttagacatca	tttatatgca	960
atttcatcag	caccagaaaag	tttgggatgt	ttttcagatg	agtaaaggac	cagggtgaaga	1020
tgttgacctt	tttgatatga	aacaatttaa	aaattcgttc	aagaaaattc	ttcagagagc	1080
attaaaaaat	gtgacagtca	gcttcagaga	aactgaggag	aatgcagtct	ggattcgaat	1140
tgccctgggg	acacagtaca	caaagccaaa	ccagtacaaa	cctacctacg	tggtgtacta	1200
ctcccagact	cgttacgcct	tcacgtcctc	ctccatgctg	aggcgcaata	caccgcttct	1260
gggtcaggag	ttagaagcta	ctgggaaaaat	ctacctccga	caagaggaga	tcatttttaga	1320
tattaccgaa	atgaagaaag	cttgcaatta	gtgaacatga	aaggaaaata	aaaattcctc	1380
acagtcaaaa	aaaaaaaaaa	aaaaa	1405			

<210> 266

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 266

ccgacaagag gagatcattt tagatattac cgaaatgaag aaagcttgca attagtgaac 60

<210> 267

<211> 927

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 267

```

ggcagcgggc gaaaggagcc ggggcctgga ggtttgcgta ccggtcgcct ggtcccggca 60
ccagcgccgc ccagtgtggt ttcccataag gaagctcttc ttctgcttg gcttcacct 120
ttaacccttc cacctgggag cgtcctctaa cacattcaga ctacaagtcc agaccagga 180
gagcaaggcc cagaaagagg tcaaaatggg gtttatattt tcaaaatcta tgaatgaaag 240
catgaaaat caaaaggagt tcatgcttat gaatgctoga cttcagctgg aaaggcagct 300
catcatgcag agtgaaatga gggaaagaca aatggccatg cggattgcgt ggtctcggga 360
attcctcaaa tatttttgga ctttttttgg ccttgcagcc atctctttaa cagctggagc 420
gattaaaaaa aagaagccag ccttctggt cccgattgtt ccattaagct ttatcctcac 480
ctaccagtat gacttgggct atggaaccct tttagaaaga atgaaaggct aagctgagga 540
catactggaa acagaaaaga gtaaatgca gctgccaaaga ggaatgatca cttttgaaag 600
cattgaaaaa gccagaaagg aacagagtag attcttcata gacaaatgaa atcatgctta 660
ccaatcaaat ctcaaagcac agaattattg acttgaatca tgggttttac agttttttaa 720
atgctcaaga ttttgatatt atagatttta ttttaaaata ttaaaatgca agatagtttt 780
gagctat tttt aaaataaaat ttataacatt caacacaaaa tcatggagggt gctctaaata 840
acttttagat ttctctcttc tgtgtgcatt accaatatct aagtgtaaaa ttaataaatt 900
gttttgaatt cctggaaaaa aaaaaaa 927

```

<210> 268

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 268

```

ggaacagagt agattcttca tagacaaatg aaatcatgct taccaatcaa atctcaaagc 60

```

<210> 269

<211> 1047

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 269

```

cccactcttc cagccagcgc cccagccctc ccgcccggcg ctgcaggtc cagaggagcg 60
cagactgtgt ccctgacaat gggaacagcc gacagtgat agatggccc ggaggcccca 120
cagcacacc acatcgatgt gcacatccac caggagtctg ccctggcca gctcctgctc 180
acctgctgct ctgcgtgctg gccccgggccc acccaggcca ggggcagca cgggtgctg 240
gtggcctcgt ggggtgatgca gatcgtgctg gggatcttga gtgcagtcct aggaggattt 300
ttctacatcc gcgactacac cctcctcgtc acctcgggag ctgccatctg gacaggggct 360
gtggctgtgc tggctggagc tgctgccttc atttacgaga aacgggggtg tacatactgg 420
gcctgctga ggactctgct aacgctggca gctttctcca cagccatcg tgcctcaaa 480
ctttggaatg aagatttccg atatggctac tcttattaca acagtgcctg ccgcatctcc 540
agctcgagtg actggaacac tccagcccc actcagagtc cagaagaagt cagaaggcta 600
cacctatgta cctccttcat ggacatgctg aaggccttgt tcagaacct tcaggccatg 660
ctcttgggtg tctggattct gctgcttctg gcactcttga cccctctgtg gctgtactgc 720
tgagaaatgt tcccaaccaa agggaaaaga gaccagaagg aaatgttgga agtgagtgga 780
atctagccat gcctctcctg attattagt cctgggtgctt ctgcaccggg cgtccctgca 840
tctgactgct ggaagaagaa ccagactgag gaaaagaggc tcttcaacag cccagttat 900
cctggcccca tgaccgtggc cacagccctg ctccagcagc acttgcccat tccttacacc 960

```


ccttccccat cctgctccgc ttcatgtccc ctccctgagta gtcattgtgat aataaaactct 1020
catgttatctg ttcccaggaa aaaaaaa 1047

<210> 270

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 270

aaccaaagggg aaaagagacc agaaggaaat gttggaagtg agtggaaatct agccatgcct 60

<210> 271

<211> 2280

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 271

ccgcccgc	ccagctacgc	cccgtccgac	gtgcctcgg	gggtcgcgct	gttcctcacc	60
atcccttctg	ccttcttctc	gcccagagctg	atatttgggt	tcttggctctg	gaccatggta	120
gcccgcaccc	acatagtata	ccccttgctg	caaggatggg	tgatgta.tgt	ctcgctcacc	180
tcgtttctca	tctccttgat	gttcctgttg	tcttacttgt	ttggatttta	caaaagattt	240
gaatcctgga	gagttctgga	cagcctgtac	cacgggacca	ctggcatcct	gtacatgagc	300
gctgccgtcc	tacaagtaca	tgccacgatt	gtttctgaga	aactgctgga	cccaagaatt	360
tactacattta	attcggcagc	ctcgttcttc	gccttcatcg	ccacgctgct	ctacattctc	420
catgccttca	gcattctatta	ccactgatgc	acaggcgcca	ggccaagggg	gaaatgctct	480
ttgaaagctc	caattattgg	tcccctaaaag	cagcttccaa	cgtttgc cat	ctggatgaca	540
aacggaagat	ccactaaaac	gtccacggga	ttaacagaac	gtccttgcag	actgagcgat	600
gacaccacac	tttgtttgga	cattttaaatt	cactctgctg	aatagga.gga	agcttttctt	660
tttcctggga	aaacaactgt	ctcttggaat	tatctgacca	tgaactt gct	cttctagaca	720
actcacatca	aagccctcac	tccactaatg	gagaatccta	gccccac taa	tgccaagtct	780
gtttggggat	tttgcctcag	ctatgggctt	ccctagagta	ggctctagggg	aatactcagt	840
ctgatcttct	ttttgtttgt	tttattttgt	tttttttgag	acggagtctc	gctcttctct	900
caaggctgga	gtgcagtga	gcgatctcca	ctcactgcag	gctccgcctc	ccgggttccc	960
gccattctcc	tgctcagcc	tcccagagtag	ccgggactac	aggcgcc cac	caccatgccc	1020
ggctaattta	gttgtatttt	tagtagagat	ggggtttcac	cgtattagcc	aggatggtct	1080
cgatctcttg	acctcgtgat	ccgcccgcct	cgccctccca	aagtgcggg	attacaggcg	1140
tgagccaccg	tgcccggcct	gattctctta	aaattgaaga	ggtgctgcca	aggccttcag	1200
atctaaccga	gatgcataga	ccttgttcct	ggtacttgtt	cagcctgtgc	tggggagccg	1260
tggtcccrag	ttccctggga	ggctgacagg	gtcaagccac	cctgcccacc	accctcccac	1320
ttcccctccc	ctttcctctc	cagcattagg	attcaaggga	aatctgcatg	aagccaattt	1380
tgagggtaga	cgtgtgggga	aaataaatca	ttatacagta	agacctgggg	cttgaggggt	1440
ggggaatggg	gagggaaggg	catagcctgc	tccctcatga	gtctgacatc	tcggaaactg	1500
agcagctgcc	ggacgcctgg	gtcaggaatc	caagacccca	cctcttaagg	actgggtcct	1560
cagaaagcac	cctcagggaa	aaagggtgaaa	acattacatc	cgtggat tct	cctgccacaa	1620
ccgcattgga	agaaaaaggct	gccgcaacat	ctcagcgagg	agtgaaggac	ccatgtccca	1680
ggaaccgcgc	tgcgccacct	gcactcacc	ccctcacatt	ctcttaagca	ccgggtggcc	1740
ctccagggct	ggcggaatgg	tggtgcccac	ggggttgggc	aagggtcac	caggagctca	1800
acgggcaaag	ttgtgcacac	taaaatatca	aatcaagggtg	cttgggtt tta	aagtaaattgt	1860
ttttctaaag	aaagctgtgt	tcttctgttg	accagacga	atagggcaca	gcctgtaac	1920
tgcacgtgcc	ttctgtcatt	gggaatgaaa	taaattatta	cgagaaaagg	acttgtccta	1980
actggtttga	ggccttacag	ttttgtatct	acatttttcc	cctcctgggg	tttgcgggga	2040
cagggacaga	actacaggag	tcatgggaaa	gaaaattctg	gcttcactac	tgctcactgc	2100
tcactttctg	atcactctga	tacttttttt	tttttttttt	ttttgcaacc	tgataccttg	2160
aaaagcttct	atgtgtctct	ccttttggtg	cctggcagct	gtctaggatg	atcactgatt	2220

actattttact aagtagccac atgcaaataa aagttgtttg gtaaaatgga aaaaaaaaaa 2280

<210> 272

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 272

tcaccaggac ctcaacgggc aaagttgtgc acactaaaat atcaaatcaa ggtgcttggg 60

<210> 273

<211> 2554

<212> DNA

<213> Homo sapiens

<300>

<308> X87949

<400> 273

aggtcgacgc	cggccaaagac	agcacagaca	gattgaccta	ttgggggtggt	tcgcgagtggt	60
gagaggggaag	cgccgcggcc	tgtattttcta	gacctgccct	tcgcctgggt	cgtgggcgct	120
tgtgaccctcg	ggcccctgcc	gcctgcaagt	cggaaattgc	gctgtgctcc	tgtgctacgg	180
cctgtggctg	gactgcctgc	tgctgcccaa	ctggctggca	agatgaaagt	ctccctgggt	240
gcgcgatgc	tgctgctgct	cagcgcggcg	cgggccgagg	aggaggacaa	gaaggaggac	300
gtgggcaagg	tggtcggcat	cgacttgggg	accacctact	cctgcgtcgg	cgtgttcaag	360
aacggccgcg	tggagatcat	cgccaacgat	cagggcaacc	gcacacagcc	gtcctatgtc	420
gccttcactc	ctgaagggga	acgtctgatt	ggcgatgccg	ccaagaaaca	gctcacctcc	480
aaccccgaga	acacgggtctt	tgacgccaaag	cggctcatcg	gccgcacgtg	gaatgacccg	540
tctgtgcagc	aggacatcaa	gttcttggccg	ttcaaggtgg	ttgaaaaaga	aactaaacca	600
tacattcaag	ttgataattgg	aggtggggcaa	acaaagacat	ttgctcctga	agaaatttct	660
gccattggttc	tcactaaaat	gaaagaaacc	gctgaggctt	atttgggaaa	gaaggttacc	720
catgcagtgt	ttactgtacc	agcctatctt	aatgatgcc	aacgccaagc	aaccaaagac	780
gctggaaacta	ttgctggcct	aaatgttatg	aggatcatca	acgagcctac	ggcagctgct	840
attgcttatg	gcctggataa	gagggagggg	gagaagaaca	tcctgggtgt	tgacctgggt	900
ggcggaacct	tcgatgtgtc	tcttctcacc	attgacaatg	gtgtcttcga	agttgtggcc	960
actaatggag	atactcatct	gggtggagaa	gactttgacc	agcgtgtcat	ggaacacttc	1020
atcaaaactgt	acaaaaagaa	gacgggcaaa	gatgtcagga	aggacaaatag	agctgtgcag	1080
aaactccggc	gcgaggtaga	aaaggccaag	gccctgtctt	ctcagcatca	agcaagaatt	1140
gaaattgagt	ccttctatga	aggagaagac	ttttctgaga	ccctgactcg	ggccaaattt	1200
gaagagctca	acatggatct	gttccggctc	actatgaagc	ccgtccagaa	agtgttggaa	1260
gattctgatt	tgaagaagtc	tgatattgat	gaaattgttc	ttgttgggtg	ctcgactcga	1320
attccaaaga	ttcagcaact	ggttaaagag	ttcttcaatg	gcaaggaaac	atcccggtgc	1380
ataaaccag	atgaagctgt	agcgtatggg	gctgctgtcc	aggctgggtgt	gctctctggt	1440
gatcaagata	caggtgacct	ggtactgctt	catgtatgtc	cccttacact	tggtattgaa	1500
actgtaggag	gtgtcatgac	caaactgatt	ccaagtaata	cagtgggtgc	taccaagaac	1560
tctcagatct	tttctacagc	ttctgataat	caaccaactg	ttacaatcaa	ggtctatgaa	1620
ggtgaaagac	ccctgacaaa	agacaatcat	cttctgggta	catttgatct	gactggaatt	1680
cctcctgctc	ctcgtggggg	cccacagatt	gaagtcacct	ttgagataga	tgtgaatggg	1740
attctctcag	tgacagctga	agacaagggg	acagggaaca	aaaataagat	cacaatcacc	1800
aatgaccaga	atcgctgtac	acctgaagaa	atcgaaagga	tggttaattga	tgctgagaag	1860
tttgctgagg	aagacaaaaa	gctcaaggag	cgcatgtata	ctagaaatga	gttggaagac	1920
tatgcctatt	ctctaaagaa	tcagattgga	gataaagaaa	agctgggagg	taaactttcc	1980
tctgaagata	aggagaccat	ggaaaaagct	gtagaagaaa	agattgaatg	gctggaaagc	2040
caccaagatg	ctgacattga	agacttcaaa	gctaagaaga	aggaaactga	agaaattggt	2100
caaccaattta	tcagcaaaact	ctatggaagt	gcaggccctc	ccccaaactg	tgaagaggat	2160
acagcagaaa	aagatgagtt	gtagacactg	atctgctagt	gctgtaatat	tgtaaatact	2220
ggactcaggga	actttttgta	ggaaaaaatt	gaaagaactt	aagtctcgaa	tgtaatggga	2280
atcttcacct	cagagtggag	ttgaactgct	atagcctaag	cggctgttta	ctgcttttca	2340

```

ttagcagttg ctcacatgtc tttgggtggg gggggagaag aagaattggc catcttaaaa 2400
agcgggtaaa aaacctgggt taggggtgtgt gttcaccttc aaaatgttct atttaacaac 2460
tgggtcatgt gcatctgggt taggaagttt tttctacat aagtgcaccc aataaatgtt 2520
tgttatttac actggtcaaa aaaaaaaaaa aaaa 2554

```

<210> 274
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> X87949

<400> 274
 aactttcctc tgaagataag gagaccatgg aaaaagctgt agaagaaaag attgaatggc 60

<210> 275
 <211> 1359
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 275
 ttttaagaca gttacctgtt gtgctgctgt tacaatatat aatgraaacca agtcagggga 60
 gtgaatttat caatcttttg atgtaaagta aaaacgtagt tcacacttca ggagagaact 120
 tcatagcaca atgtctttct ataagatatt tttaatgatt tagtatttta caacatttgt 180
 ttaccatatt ttgatatacc atttttttct atctgcccag ttttattaaa aaaactatat 240
 attattttct aaagaaacaa tcatattttt atacaaaatt atgttttcag gtaacgaaat 300
 agatgtaggg tacagtggaa cataagcagt gttaccctg gctgggagtc agtattatac 360
 aacaaatggt gagctggaac atgccctgtc tgtgctgtcc ctcc tgtgct gggtcgcgga 420
 tgtgtaggca acattgcctt atcacgctag gttcacctga cact ttaaaa ggaaaaaaag 480
 ttccatagag ttctgtggctc acaaaattgt tttgctttta tcaaatactt taatagaacc 540
 aaagttgcag atattggaat gtatggaagt atctcagtct ctgcataaga ggattaaagt 600
 atgaaaggat catttaatga ctgttttact tataagtcac taagtaatcc accatttctt 660
 atggatgatg ctttaagcctg gtgagggttg tactctaagg agccagatc ataatgcagt 720
 gcatttcctt agcccttaga gtttcttgca aacattttaa aaaaagacata ttttaagaaag 780
 aaagataaag aaaaaacata ttttaattact gtaaacaggt actgctttat gtttattttc 840
 tctctacttc aaccaaaatc agatctttga ggttttgctg acattgttgg tggttttgca 900
 catgttcttt ctaattggat ttatgaatag ttctatgggt tttaaagat gaatcatgct 960
 aagaacactt ctgctttttg atccactgtt tgcagcagaa ttatatatat gtataggaaa 1020
 aatccacttt gaataatcca tgttttgtat ttggaaattg tttttaaaaa taaaaaggaa 1080
 aggaaatata taaagctgtt atttattctg catttcttac atatctatcg cttgtcagta 1140
 taccgctttt ggtatatatt gcctctgcac atctacattt gtatatgcaa cagtgcagct 1200
 tatatctaca taaactgtaa ataatccttt ctgtgaaagg atcatcatat caagatgata 1260
 ccaaaagtat gtaaaaagaa acctgcatta ttttgtaatt atttcttata gatatttcat 1320
 ggtaagatta gcagtcaata aagttacttt tttgccttt 1359

<210> 276
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 276
 gggttttcaa agatgaatca tgctaagaac acttctgctt tttgatccac tgtttgcagc 60

<210> 277

<211> 994
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 277
 tgaatgtata tattaagact gtagctgaat tgcacatgaa atcagattgc caacttcttg 60
 actttcaatg ttagacattt atccttaagt tgtgagcgat atatgtagca tgctgtgaaa 120
 tgtctgttat agctctttta ttcatcagta ttaatacaga attatcattt gcgtttcttg 180
 gtacttttta ttcaatgtaa tcagaagctg tgatgttttg cctttgtagt cctgtgcttt 240
 gttactgtaa tttttttttt tttttttacg aagcacgtga ctgga.ctaat gtaaggcaga 300
 tgacgtgata ttttaagactg ctatatatat cagtctctta ctcta.taagg ttttaaatta 360
 gaataagctt ttatcaaata gataattgat gcaatttagg attca.cgcaa gtttcagtgt 420
 caaatggcgg tcttatagtt tcaattctga aaatagcaaa cttaa.taaac agccacttta 480
 aacttgttct ggcaaaccag accctgctgt agatatagtc taaggtagtt aaccatataa 540
 gccttttcaa ctcttaatgc cctccacatg aatcagcagt taaga.aggtt ctagaaccca 600
 tgaaagcttt tgtatgtatt actagggttt gttttcttta tgttt.gctga ttttacagtt 660
 ctgactaaag ctgacctaaa tggatcagtt tatgtgtaat attct.agtgc tttaatgact 720
 ctttttttct ttggagggag ggtaacatta tttggacaga tgcagaagga actgttagtg 780
 agtcaagaca aacacatctg aaataaagga actgtgtatt aacat.gttta caattcataa 840
 ctgcaacttt tatgacattt tgaaaatcta tttataggta cagaa.caatg ggttttgtta 900
 aactgtatca catttatact tgcagaaaatt tatttcattg ttatt.agtag gaattttatt 960
 ggttcaataa aattggcaaa actgaacacc aaaa 994

<210> 278
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 278
 ctgctgtaga tatagtctaa ggtagttaac catataagcc ttttcaactc ttaatgccct 60

<210> 279
 <211> 423
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

<400> 279
 tatgttatgg atatcttatt ttagagtaag aatataaggc atagc.catat ttatgaaggt 60
 agtaatactc tactaatcaa tacttagaag tttttgttat gacta.atctg aatgcttttt 120
 agttttt.cct taatctagtt atgttggtta tttataagtc agtttt.caga ttaggaaaga 180
 aggtattttga ggggtgttcca tttccactga atagtaagat gatgc.ttact tagattttcca 240
 cagctgtttg aaagctctgt atttggctat aacggaaaac tttgt.taggg atgcttgatg 300
 ttttgtgttt tgtttctaaa ggaagacagt gttttgttcc ttctt.tagaa aacttgaaga 360
 atagaataat gagtccagga ttaatttggg ataaagtctt ttact.tcata aattctgatt 420

ctg 423

<210> 280
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> Contig14683

<400> 280

aggaagacag tgtttttgttc cttcttttaga aaacttgaag aatagaataa tgagtccagg 60

<210> 281

<211> 391

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28552

<400> 281

atgccattga	tgtgaagaag	gtgtctgtgg	aagactttct	tactgacctg	aataacttca	60
gaaccacatt	catgcaagca	ataaaggaga	atatcaaaaa	aagagaagca	gaggaaaaag	120
aaaaacgtgt	cagaatagct	aaagaattag	cagagcgaga	aagactcgaa	cgccaacaaa	180
agaaaaagcg	tttattagaa	atgaagactg	aggggtgatga	gacaggagtg	atggataatc	240
tgctggaggc	cttgcaagtc	ggggctgcct	tccgcgacag	aagaaaaagg	acaccgatgc	300
caaaagatgt	tcggcagagt	ctcagtcctc	tgtctcagag	gcctgttctg	aaagtttgta	360
accatggtta	taaaccgtat	ttataaattg	c			391

<210> 282

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28552

<400> 282

aagactttct tactgacctg aataacttca gaaccacatt catgcaagca ataaaggaga 60

<210> 283

<211> 450

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 283

ctcatccaag	gagctggggc	agacttcatt	gattctagag	agacctgttt	cagtgcctac	60
tcatccctgc	cctctgggtg	cagcctcctt	accatcacgg	cttcaactgag	gtgtagggtgg	120
gtttttctta	aacaggagac	agtctctccc	ctcttacctc	aacttcttgg	gggtgggaatc	180
agtgatactg	gagatggcta	gttgctgtgt	tacgggtttg	agttacattt	ggctataaaa	240
caatcttggt	gggaaaaatg	tgggggagag	gacttcttcc	tacacgcgca	ttgagacaga	300
ttccaactgg	ttaatgatat	tgtttgtaag	aaagagattc	tgttgggtga	ctgcctaaag	360
agaaagggtg	gatggccttc	agattatacc	agcttagcta	gcattactaa	ccaactgatg	420
gaagctctga	aaataaaaaga	tcttgaaccc				450

<210> 284

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 284

agacagattc caactgggta atgatattgt ttgtaagaaa gagattctgt tggttgactg 60

<210> 285

<211> 439

<212> DNA

<213> Homo sapiens

<300>

<308> Contig30875

<400> 285

agaaatcaat gacagttgac aggaagagag gacgcataca acaggcaaaa gaggaatgcc 60
 cagcagtcctt ggtccttgcg gtgcaatact ggccttgagg ccaagtcagc aggggattcg 120
 tagtcactaa cttctaactg aggcagggaa gtaccatgtt ctggaaaagg tccaaagaaa 180
 caggaataga ggcagtgtag caagaggcag atttttgggtg ccaaatagat ttgaatcctg 240
 gtctctgcttc ttcctttgta gagtatgata ttggttcctt cctcccaaag ctattataaa 300
 gactaaatat gtacacaaat ctttgggatg tctgacatat aaatgcttaa caatagggtat 360
 ttgctgggtat tattacaaat gaatttgctt atttttgagc cacttctatg tctgtccatt 420
 aaaccaaatt gtgttctgc 439

<210> 286

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig30875

<400> 286

ggttctttcc tcccaaagct attataaaga ctaaatatgt acacaaatct ttgggatgtc 60

<210> 287

<211> 338

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31221

<400> 287

gggaagttac actgcttcac accacaaggc cgtgggaaat cttggagggtt ctgtgccttt 60
 ctgtcacctc tactttttgc agctgtgatt gcactgtccc gcaatgtga ctacaagcat 120
 cactggcaag gaccctttaa atgggtgaaaa tgggcagatg aatagcaata agtggacctt 180
 tgttactctt ctgagttaga aaaattctaa tttagtacac tctgaacaaa gcttattata 240
 ct tacttaag atgtgttttg atttgggtgt cagaaagcaa cctgacaatg ataatactgt 300
 aaatatgata aaattgagaa taaaagatt ttatttag 338

<210> 288

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig31221

<400> 288

aaatgggcag atgaatagca ataagtggac ctttgttact cttctgagtt agaaaaattc 60

<210> 289

<211> 417

<212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31288

<400> 289
 gaatcacttg agcccgggag gttgaggctg cagtgaagctg tgtttataacc actgcactcc 60
 agcctgctgg gtaacagagc aagactccat ctcaaaaaga aaagaaaaaa tgctttgcta 120
 cataatgagg ccaggcaaaa aaaaaaaaaaag tctgttgga atcatataga caaacatttg 180
 caaagctgct actgccattg taccagtgtt aaaatgtgtt ctaccttgca tcttttactg 240
 atttttatga cagattttat attgtaacca tttgagaact ctgtaagtgc tatggcttcc 300
 ttaaaactacg atttatcata tgctcccagt gtttactttg agactgaatg gcaaccagag 360
 aatgtaaaca accaaggtgc atctggttat gttttaaaat aaagattaat aaaagtt 417

<210> 290
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31288

<400> 290
 ggcttcctta aactacgatt tatcatatgc tcccagtgtt tactttgaga ctgaatggca 60

<210> 291
 <211> 394
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31646

<400> 291
 gctgctacac cccatgtaaa aagcggaaaa taaaatgaag attttccagc gcaagatgcg 60
 gtactgggttg cttccacctt ttttggaat tgtttatttc tgcaccattg tccaaggtca 120
 agtgggtcca cccacaaggt taagatataa tgtaatatct catgacagta tacagatttc 180
 atggaagggt ccaagaggga aatttggttg ttacaaactt cttgtgactc caacttcagg 240
 tggaaaaact aaccagctga atctgcagaa cactgcaact aaagcaatta ttcaaggcct 300
 tatgccagac cagaattaca cagttcaa atttgcatac aataaagata aagaaagcaa 360
 gccagctcaa ggccaattca gaattaaaga ttta 394

<210> 292
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig31646

<400> 292
 gccagaccag aattacacag ttcaaattat tgcatacaat aaagataaag aaagcaagcc 60

<210> 293
 <211> 357
 <212> DNA
 <213> Homo sapiens

<300>

<308> Contig37562

<400> 293

```
caattatttc aagtgcacct tattaacaaa agtatcagtg gatccaacat aaaattttat 60
agtactaaat gtcaagccta actgtgaatt ttgttctgta tcttaagtaa atttatgata 120
atgttctcga gctatcaaca aaatatatgt acttttgtga gctatgaatt ttctaattaa 180
attttacatg ctataacatg atttttacat gaatgatact ttgtttataa ctatcaaagtg 240
tcagtatttt actacaattt tattataaag tgtacattat cactaaatga acttcgattt 300
taaaaatcaa attagcttta gttgtatatt attttttaca aataaagata gacttgt 357
```

<210> 294

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37562

<400> 294

```
atcaaagtgc agtatttttac tacaatttta ttataaagtg tacattatca ctaaataaac 60
```

<210> 295

<211> 351

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37895

<400> 295

```
aataagagaca cctctaatta attaaagcgg atgccctccc cactcctccc aggatttgac 60
tcggagcaca aactcttcac aaaccaaaat gtcaggacac catcgccagt gtccactggc 120
cactgctggt ggtgtgaggc agccaggagc cctcagaac tagtaagtct gagaagaggc 180
tgcacggggc ctaggagagg gagaaatgag cccgtccaag gtgaattcct tgattctcca 240
ttgtgagtg accaagaaca agcactccct ccgactgact ctccgctacc aggatctgga 300
acaCcttcca ttaatttatt cgttcattca ataaatattt attgactgac t 351
```

<210> 296

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig37895

<400> 296

```
ctctcgcta ccaggatctg gaacaccttc cattaattta ttcggttcatt caataaatat 60
```

<210> 297

<211> 418

<212> DNA

<213> Homo sapiens

<300>

<308> Contig38288

<400> 297

```
gacaagtaaa tgggggccgt tgggacggcg ggtgcctgga gggcagctct gggctcagcg 60
ggcagtgcct agagcacagg cccctctgtt gggggatggg gaggagagca gtctgccctt 120
gggaCcgtag gccccaggga gacttctaaa gccccctgt tcgtctgctc ttcacccagc 180
accaCagagg cacctgctgc acacacaagc atctcactcg gccacaggag ggggccaggc 240
```


ttcctttgcc tgaagctggt ttgggaaggg tctccacaca ggcactgac tcccaagctt 300
 tggatcatgat gtcttttacc atttgataat tttaaacatt gtttttaaac ccaaaacatt 360
 tagtggtccg ttgcctctga agatgtaaac aaacaaatac actatttctg ggaacatt 418

<210> 298
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38288

<400> 298
 ttttagtggtc cggtgcctct gaagatgtaa acaaacaat acactatttc tgggaacatt 60

<210> 299
 <211> 413
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 299
 tacatttttg tttaatgttg ggcctgaggt taactgtgac catggtccag cttgagtggc 60
 ttctggagca gccacatttt caaggactgt ccaaaagcca gccagttcag ggctcaggcc 120
 tccccatttg cccactcctg gggagaccat cacctggctc atcgtttcca ccaagagtgc 180
 cccacaggag tgccccacag acccgctgga ccagcctgct gcgggtcctg gccaggggtc 240
 tggctaacgg tgagggctga ctctgaactg tctctcagtc tccagaaagt gttcaagcct 300
 gtgtgtgtcc caaatctgat tctcctatt gtcttgtaa tcaaactcta agtgaaaact 360
 tccccatttg cccttcaaag attttttttt attaaatggg tttttaagat cct 413

<210> 300
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 300
 tgttcccaaa tctgattcct cctattgtct tgtaaatac actctaagtg aaaacttccc 60

<210> 301
 <211> 434
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 301
 gaatgggtgaa agagagatgc cgtgttttga aagtaagatg atgaaatgaa tttttaattc 60
 aagaaacatt cagaaacata ggaattaaaa cttagagaaa tgatctaatt tccctgttca 120
 cacaactttt acactttaat ctgatgattg gatattttat ttttagtgaac catcatcttg 180
 ttagctaact ttaaaaaatg gatgtagaat gattaaagggt tgggtatgatt tttttttaat 240
 gtatcagttt gaacctagaa tattgaatta aaatgctgtc tcagtatttt aaaagcaaaa 300
 aagggaatgga ggaaaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 360
 ctagaaatga gataaagatt atttattttt gttcatatct tgtacttttc tattaaaatc 420
 atttttatgaa atcc 434

<210> 302
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 302
 aaggaatgga ggaaaattgc atcttagacc atttttata t gcagtgtaca atttgctggg 60

<210> 303
 <211> 391
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40552

<400> 303
 caccaagccc tgctccggca cctogaatcc ctggcgacca tgagtcacca gctccaagcc 60
 ttactgtgcc ccagaccaa gagctccatc ccccgccctc tgcagcgttt gtctagcgcc 120
 cttgcagctc cagagccccc tggcccagcc cgtgactcct ctttggggcc tacagatgaa 180
 gctggctctg agtgtccctt ccctagaaag gcctgaccct ccttaccac cagaacaggg 240
 gttttgatgc cctcactagt gttgaagcct gttccagaga gaggtgggac tgcaaggaga 300
 ggatggtcag ccctaccac ctgccctgtt tgagcttcc ttttgacaat gtttgctgtt 360
 gattttttgt tcaataaaga atttggtaaa a 391

<210> 304
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40552

<400> 304
 tttgagcttc ctgtttgaca atgtttgctg ttgatttttt gttcaataaa gaatttggtg 60

<210> 305
 <211> 495
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig41413

<400> 305
 aaatattctt aatagggcta ctttgaatta atctgccttt atgtttggga gaagaaagct 60
 gagacattgc atgaaagatg atgagagata aatgttgatc ttttggcccc atttgттаат 120
 tgtattcagt atttgaacgt cgtcctgttt gttgtagtt ttcttcatca tttattgtat 180
 agacaatttt taaatctctg taatatgata cattttccta tcttttaagt tattgttacc 240
 taaagttaat ccagattata tggtccttat atgtgtacaa cattaaaatg aaaggctttg 300
 tcttgcattg tgaggtagag gcggaagttg gaatcagggt ttaggattct gtctctcatt 360
 agctgaataa tgtgaggatt aacttctgcc agctcagacc atttcctaат cagttgaaag 420
 ggaaacaagt atttcagtct caaaattgaa таатgcacaa gtcttaagtг attaaaataa 480
 aactgttctt atgtc 495

<210> 306
 <211> 60
 <212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

<400> 306

cagctcagac catttcctaa tcagttgaaa gggaaacaag tatttcagtc tcaaaattga 60

<210> 307

<211> 409

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 307

aaaaaaaaa aaaaaaaaaa aaagagttgt tttctcatgt tcattatagt tcattacagt 60
 tacatagtcc gaaggtctta caactaatca ctggtagcaa taaatgcttc aggccacat 120
 gatgctgatt agttctcagt ttctattcag ttcacaatat aaccaccatt cctgccctcc 180
 ctgccaaggg tcataaatgg tgactgccta acaacaaaat ttgcagtctc atctcatttt 240
 catccagact tctggaactc aaagattaac ttttgactaa ccctggaata tctcttatct 300
 cacttatatgc ttcaggcatg tatttatatg tattcttgat agcaatacca taatcaatgt 360
 gtattcctga tagtaatgct acaataaatc caaacatttc aactctgtt 409

<210> 308

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 308

ctcatgttca ttatagttca ttacagttac atagtccgaa ggtcttacia ctaatcactg 60

<210> 309

<211> 552

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 309

ctgaagacta cgaccatgaa atcacagggc tgcgggtgtc tgtaggtctt ctcttggtga 60
 aaagtgtcca ggtgaaactt ggagactcct gggacgtgaa actgggagcc ttaggtggga 120
 ataccagga agtcaccctg cagccaggcg aatacatcac aaaagtcttt gtcgccttcc 180
 aagctttcct ccgggggtatg gtcatgtaca ccagcaagga ccgctatttc tattttggga 240
 agcttgatgg ccagatctcc tctgcctacc ccagccaaaga ggggcagggtg ctggtgggca 300
 tctatggcca gtatcaactc cttggcatca agagcattgg ctttgaatgg aattatccac 360
 tagaggagcc gaccactgag ccaccagtta atctcacata ctcagcaaac tcaccctggg 420
 gtcgctaggg tgggggtatgg ggccatccga gctgaggcca tctgtgtggt ggtggctgat 480
 ggtactggag taactgagtc gggacgctga atctgaatcc accaataaat aaagcttctg 540
 cagaatcagt gc 552

<210> 310

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41887

<400> 310

tactggagta actgagtcgg gacgctgaat ctgaatccac caataaataa agcttctgca 60

<210> 311

<211> 745

<212> DNA

<213> Homo sapiens

<300>

<308> Contig42342

<400> 311

gcagtaaaga caggacgcac ccatgtcaca agaggagcac aggcaggggt gttggtgttg 60
 gggcagccct cagggtctcc agaccagcc cactcacac agcagcctag gaaggaagg 120
 cagagtccca ggtgtcagct ggtgggtctc ccaggagctg cccctccctg gaagtcacag 180
 gacaggaatg acagatcagg gaactgcagg aagctgccac ctctggggtc agaatatgcc 240
 cagcctgcgg gggctctcta tcgggggtctt cgagagccag acagcctgcc ttgtgctgca 300
 tacctggctt tgctctgtgc agaaccagc acacgtgatt ttgtgtgaca tgccagcagc 360
 ctggctccca ggacaggagg cctgccctgg gggaggggct gcaggaggag ggggggcagg 420
 caccatgag tctgtccagc cttgtcacag atgcatcgcc caagctgcgg tcctgatttc 480
 agctcacctc agagtaaate agaataaact gcaccagac tttcacgaat gcatgttgac 540
 gctttcagtt caccctttc tttgctaact ttcttctat tttcttctaa tgcgagagct 600
 tattaattcc atatttatca ttttgaataa cttttctcct ttttagtaac aaaatgtact 660
 tcaactcttag taaaatgtat ttactatttt agtaacaaaa atatacttgc ctaatcatgt 720
 ttaaaatata gtgatgtgaa aaatt 745

<210> 312

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig42342

<400> 312

caccagact ttcacgaatg catgttgacg ctttcagttc acccctttct ttgctaactt 60

<210> 313

<211> 398

<212> DNA

<213> Homo sapiens

<300>

<308> Contig43645

<400> 313

agttcaaagg cagataaatc tgtaaattat tttatcttat ctaccatttc ttaagaagac 60
 attactccaa aataattaaa ttttaaggctt tatcaggtct gcataatagaa tcttaaattc 120
 taataaagtt tcatgttaat gtcataaggat ttttaaaga gctataggta atttctgtat 180
 aatatgtgta tattaataatg taattgattt cagttgaaag tatttttaaag ctgataaata 240
 gcattagggg tctttgcaat gtggtatcta gctgtattat tggttttatt tactttaaac 300
 attttgaaaa gcttatactg gcagcctaga aaaacaaca attaatgtat ctttatgtcc 360
 ctggcacatg aataaacttt gctgtggttt actaatct 398

<210> 314

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig43645

<400> 314

gaaaagctta tactggcagc ctagaaaaac aaacaattaa tgtatcttta tgtccctggc 60

<210> 315

<211> 478

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44289

<400> 315

ctaaaaacaa cactcatcag tcttgggaaa tttgaacttt gatcaactta actaaagaag 60
gaagggtagt aagaattttt caaatacaaa tatttgccaa ttcacagatg ataacattta 120
aggccttcaa aagtaagggt ttttccttgt ttctccagtc agcttttgtc aactctaata 180
gttttttcat aaacattttt tatttgtata attgcaacag ttttaagaaat tatcacaact 240
atttagaaac atttaaaatg ttctttttga tataagctat atacttggaa aaatacattg 300
gtatctaaaa tttgaggtgt gttaagactg ctttttggtt taaaaaatgg tttacattca 360
aatttttgaa gtgttttatg cttcatatgg ctaagttgta gtttggcaga gttaacagca 420
taagaataaa catgctgtaa ttttaaaaga tgctt tgaat aaaaatttat ttttaattt 478

<210> 316

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44289

<400> 316

catcagtctt gggaaatttg aactttgatc aacttaacta aagaaggaag ggtagtaaga 60

<210> 317

<211> 556

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44909

<400> 317

accatctggg atttctacag cctgggtacc catagccaca ccaaggcttc tgggagattc 60
tgcagggtca gctttccagg ctgttcccaa atagctccct gcctccccac tgccccataa 120
gccacagcag aagagccatt catctcataa acaaaaagga agaggaaaga atgagggaagg 180
accctgtgca aggttatattg caggcaggga tgggc ttgta cctgacagca cccacccctg 240
tgtggccccc aggccctcat caccctcaga cccct cctaa gcagttccct cattgctctt 300
tggactaggc tgacagcagg aagagcaggg cccatgaccg ggtggaagtt cagttttggt 360
gtctgcttca agaggggggtt ttacactctg attccaggac aagcactctg aggcgggtgg 420
gggagagaaa ccctggctct tcacccagggt ttcacacaca tgtaaatgaa acactatggt 480
agtatctaac acactcctgg atacagaaca caagtcttgg cacatatgtg atggaaataa 540
agtgttttgc aatctt 556

<210> 318

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig44909

<400> 318

tcacccagggt ttcacacaca tgtaaatagaa acactatgtt agtatctaac acactcctgg 60

<210> 319

<211> 710

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 319

aaagataggc ttctaagtta aggcaaatca ttcattctgt cattaaacaa atacaaacca 60
 ggcacctgtc atatgccaag tgatattcaa aatggcccat gtagaccttt gtgaagtatg 120
 tggcctaaca gacattaaac aaatgtctgt gaaactgaca taataaagta aggtaagtta 180
 tatgtgagac attctctttt tataataatt cctgtaaagc agtacttact taggtaatga 240
 tatcactactg ttttggtttt tatttttctt aagagctaaa acgtcatcct ctcttcagtg 300
 atgtggactg ggaaaatctg cagcatcaga ctatgccttt catccccag ccagatgatg 360
 aaacagatac ctcttatttt gaagccagga atactgctca gcacctgacc gtatctggat 420
 ttagtctgta gcacaaaaat tttcctttta gtctagcctc gtgttataga atgaacttgc 480
 ataattatat actccttaat actagattga tctaaggggg aaagatcatt atttaaccta 540
 gttcaatgtg cttttaatgt acgttacagc tttcacagag ttaaaaggct gaaaggaata 600
 tagtcagtaa tttatcttaa cctcaaaact gtatatataat cttcaaagct tttttcatct 660
 atttattttg tttattgcac tttatgaaaa ctgaagcatc aataaaatta 710

<210> 320

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 320

ttaacctagt tcaatgtgct tttaatgtac gttacagctt tcacagagtt aaaaggctga 60

<210> 321

<211> 726

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 321

atacatattg ctttagagag caggtagggt gccatgtgtt cagcagtgtg tccttaagaa 60
 aataccatct ttctaagcca ctggaatttt tactttacta tttttaacat taatggatgt 120
 caggatcatca acctcaagtc tttacatatc catgtatatt ccatatatat tgtttatata 180
 ggcccaagtt tctccttaat tgggatctat atactaccag cacaacatca aaaacatgta 240
 attgaatata tcagagctat atatgtaagg aaatgactgg tgaccccat atcatcattg 300
 ttgaattcat gttaagtaga ccctctaggg gaccataagg caattgagca cataacgaaa 360
 aatgatgcaa taagaatgta tgcactctct ttgccaaatg catgtgcttt tgtgtaacgt 420
 ggatgtaaac agaattgcag tgctgccgaa attccttgatc ttggctaaga gagtatTTTT 480
 ccccttgtaa ttatgactct gagataaaat tgccattttt aaatttccaa agtaacaact 540
 ttttttattt tatgaataaa cttgggattg caatttctct gatctgacaa tcaataactt 600

taacaaagat ctaaataagt gtttcaa~~g~~ga aagttttcct aagcaaagt aatattacct 660
 catttgggca tcattactct gttaatt~~c~~ta tatcaaagga aataaacttg ctacttgcac 720
 taaatg 726

<210> 322

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 322

accataaggc aattgagcac ataacgaaa atgatgcaat aagaatgtat gcactctctt 60

<210> 323

<211> 580

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 323

ggtggtctct catccttgtg tgctgctctg ctaagagatg tccaaggcgg agccggggca 60
 agatccttcc agactcatct gtcagagccc caagcccttt agaccagag cccaaggacc 120
 atgcctttgg gacattagga ctgcagcctt tgcttctgtg tattttggag ttttggtgac 180
 ttttgtcacc tggacacact catttgttag ccatagtggg ttcccttggg cagcaacagt 240
 gcatgtacct ctggatgtca tctgagggtga gaccaccgag gccttttctc tctgtgtaca 300
 gaggggagtt aggagttgct ggactggatg cattacgagg actggggaca gggtagaggg 360
 acatccaggg atcagggcat gagtgggggc aaccccccg cctctgccct ggcatggtct 420
 ccgcatgggc tgaggtgtag ctgattggct gccacatttc ggccatgctg gctggcgtgc 480
 ccatgttgca gatatttcc cgagtccc agaatggatg gtattgaatc tcagccacat 540
 gcaacactgt gtccagcatt ctttgcaata aatacttttt 580

<210> 324

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 324

atattttccc gagttcccca gaatggatgg tattgaatct cagccacatg caacactgtg 60

<210> 325

<211> 632

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 325

gccatctagt ctgtggtttt ctgttgaa~~g~~c agtctgaatt gactaaaaca gtcacttggg 60
 gtagttataa accactttcc tgttgaaa~~g~~c agaacatgct gattcaactg ttttgttcaa 120
 tagcaatgat agattttgtt taagtccc~~c~~t acactttctt atttctaaat gatcaagagt 180
 acacttcttg gcagtgatta aggagtgt~~g~~t atctaacaga aaaaatatat ataccctgtg 240
 aacccgaata tggaattcag attgtttc~~t~~g ccctcagtat catacttaaa aaacaagcat 300
 acaaacaac ataaggggaac aaacagca~~a~~c cataacaaa acaaaccttt aaaggtgggt 360

```

ttttgctgtg ataaatgaat acggtaactct gaaggagaaa aaagttttctc aaatgagctt 4 20
aaactgcaag tgatttaaaa attagagaat ataattctta aagctattga aagtttcaac 4 80
cagaaaacct caagtgaatt ttgtatgtaa atgaaatctt gaatgtaagt tctgtgattc 5 40
ttaagcaaaa caattagctg aaaacttggt attgttgtag tttatgtagt aagtgacttg 6 00
gcacccatca gaaaataaag ggcattaaat tg 632

```

<210> 326

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 326

```

agcaacaat tagctgaaaa cttggtattg ttgtagttta tgtagtaagt gacttggcac 6 0

```

<210> 327

<211> 540

<212> DNA

<213> Homo sapiens

<300>

<308> Contig48913

<400> 327

```

accagagggg gtcctttttc cacagttaat ggatcggctg gtgtgccttc agggaggaag 6 0
agggaggtgg tcaagcttga aaaactggct ttaggatggg tctgactttg ttctccctcc 1 20
ccaagtgttc tcaacctcca ttctgcagtg ttcagagttt tagggaaagg gtttgggtgc 1 80
cccagcatcc aggtgtttgt tggcttagcg catgtgaagt gaaaaccttc tgggggttgtt 2 40
tggaagcagc tttctgggtc ttgtgattgt atcctgaggt ccagaaacct tattctccca 3 00
cgaggatcct cagtgaccat ggtggccaca cgccctggcca gcctgctggc tcctgggtga 3 60
gctgaagaac cttgcctgtg gcactt ttcg aggggtgagct ggaaccgaga gaacatggtc 4 20
cccgtgctgg gactcatgcg ggtcat ttcc tgccggcctg gtttcgcctg gtcgtgtctt 4 80
tatgagcacc atgtaagcct ccttgtattg agataattgg gcattaaaca ttaaactgca 5 40

```

<210> 328

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig48913

<400> 328

```

tatgagcacc atgtaagcct ccttgtattg agataattgg gcattaaaca ttaaactgca 6 0

```

<210> 329

<211> 534

<212> DNA

<213> Homo sapiens

<300>

<308> Contig49169

<400> 329

```

cctaattgta acatttttta aaatacatat ttgggactct tattatcaag gttctaccta 6 0
tgtaatttta caattcatgt ttcaagacat ttgccaaatg tattaccgat gcctctgaaa 1 20
aggggggtcac tgggtctcat agactgatat gaagtgcaca tatttatagt gcttagagac 1 80
caaactaatg gaaggcagac tattta cagc ttagtatatg tgacttaag tctatgtgaa 2 40
cagagaaatg cctcccgtag tgtttgaaag cgtaaagctg ataatgtaat taacaactgc 3 00

```



```

tgagagatca aagattcaac ttgccataca cctcaaattc ggagaaacag ttaatttggg 360
caaactctaca gttctgtttt tgctactcta ttgtcattcc tgtttaatac tcaactgtact 420
tgtattttgag acaaataggt gatactgaat tttatactgt tttctacttt tccattaaaa 480
cattggcacc tcaatgataa agaaatttaa ggtataaaat taaatgtaaa aatt 534

```

```

<210> 330
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49169

```

```

<400> 330
catacacctc aaattcggag aaacagttta tttgggcaaa tctacagttc tgtttttgct 60

```

```

<210> 331
<211> 602
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49388

```

```

<400> 331
tgtcagtggg ggggtctctg cagccaactg agactatctt gctgtgccct gagccttctt 60
aggggtttaga agaacagcat tcaa.aattcc ccgtcctgtc agtggtttgcc ttcgcacctc 120
ctcccctaaa gcagcgcggg gggcaataa gacccacccc ctccctgcag cttcacaggg 180
acgtcttctt cctcccccgc aaccacccca ggctcccctg ggaggctgca gttgtggtac 240
acgtcccccg tgctgggttg gccgtgactc gggggcgggg cgatcgggtc tcagcccctg 300
ccttccccag tctctgggtc acccgaattt tcccacccct gcttctcccc gagggaggtg 360
agctcttgag caagttggga cttgggcccgg ggcctggaag aatgattggc tgggaggccg 420
cgggaggagg gccaggaggc ccggaccagt tgggaggagt gagcaggccc cgggggaggg 480
ggatgagcgc agtttgctcg ctttctccc ctgccggccc cctccgccc caccacact 540
cgggacgtct tcattgaaga ttca.cttaca aaggaatgtt tcactaaata aaagaaaacc 600
ag 602

```

```

<210> 332
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig49388

```

```

<400> 332
cgggacgtct tcattgaaga ttca.cttaca aaggaatgtt tcactaaata aaagaaaacc 60

```

```

<210> 333
<211> 562
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig50728

```

```

<400> 333
gcgaatttgg gccccttgat cctctgatgg gagctgaaag gatgagaggt gggcatctag 60
atttagggag gctgttcagg ctttgagggt cccttacctg aacacataga aaccctggag 120
ctgtgactgt gtccatgtgt gtgtgtttgt ctgtgtgtgt tgcgggggat gggcacctgc 180

```

```

atgaatgtgg tagagaaaat ggctctgctc agaggggaaga tacgcatagc aaggcaggga 240
ccagaggaat cacaggcgcc tggagagcag ccgggcaccg cctccaggga cctgccggct 300
tccctcagtc ctccaggggc ccagcactct tccttttaggc cctgtgagcg tcccttgtca 360
ggatacattc tctcattttg ctgaagctga tttgattggg tgtctgtttc tcgcagccaa 420
aagagctctg aatgaggaaa gtgctttctgt gctaactccc cgcgtctcct gaatttca 480
cattcatgta ccgcctcga aatttttgca atatctgtgt accaactgtc catttact 540
ataaagaagt tttctttaaa tt 562

```

<210> 334

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50728

<400> 334

```

tttgattggg tgtctgtttc tcgcagccaa aagagctctg aatgaggaaa gtgctttctgt 60

```

<210> 335

<211> 400

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 335

```

tttttttttt tgcacttatg gtatttattg ttggaagatt gactacctta atgcacacca 60
atgctcagat gacttggggg cacatagggg actgctgtca ccatgcctca ctctgcagg 120
gaaggggctg cctactaaa accccagcgg gccagtgct gtgtccagaa caggtcctta 180
tattactgca gccacaatg gaactactga gtaggagcca aaagaggagg gagcaggaa 240
aggtggcatt tggagagggg agaccgcacc cacaggtctg ccacagcgcg tcaacgggat 300
ggggtacttt tacagtcaa gtgacttcgg tgtccgccca ccatctacct ttgtaggacc 360
actgaaacaa gggacatcca ccacggccca cagccggggc 400

```

<210> 336

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 336

```

gagcattggg gtgcattaag gtactcaatc ttccaacaat aaataccata agtgcaaaaa 60

```

<210> 337

<211> 475

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50950

<400> 337

```

ctggaagagg ctcccaacc agagtgtccc tgtgggaggc aggcagaagg tgacaattga 60
cacgatttcc tgcacgcgtc ctctcttacc ttggaagcag ttagaatcta ccaggcacag 120
atgaggccgc ccttgcttga cggagcttga tgagcagccc ttggtctcgg gttccaggac 180
tgagagccca gctgcctctg cccacccttc ccaggcctc tgccagcctc tggctgcacg 240
gtcaggccct gcccatggc aggcctgcca gagcttggct ggggacccct cccgcctctg 300

```

gctccctgat gggctggatg taacttgtgt cttctagccc ctttaaggagc ccaggtgttt 360
 taaggaatga attggtcact gcactcttga tcgattatgg ttctgagaaa agrcaaataac 420
 acttttggct gcattaaaag aagcatcata tataaaataa agaagatgaa ggtct 475

<210> 338
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50950

<400> 338
 gtcactgcat cttgtatcga ttatggttct gagaaaagca aatatcactt ttggctgcat 60

<210> 339
 <211> 860
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig51660

<400> 339
 ggatggcaac cttcagctag actgcctggc tcaagggtgg aagcaatacc aaacagagagc 60
 atttggctgg ttccggtggt cctcctgcc gcgaagttgg gcttccgcc agtgcagatt 120
 ctgtgccaca cgtactggga gcaactggaca tcccagggtc aggtgcgtat gaggtctctt 180
 ggccaaaggt gccagaagt ctcctgggtc caatatgaga tgcttgagtt ctctcggat 240
 agcaccatga ggattctgag caacctgggt cagcatatac tgaagaaata ctatggaaat 300
 ggcattgagga agtctcaga aatgccagta atcctggaag tgccttgga aggatcccat 360
 gacacagcca attgtgaggg atgcactttg ggcattatgtg gacagggctt aaaaagctac 420
 atgacaaagc cgtccaaatc cctactcccc cacctaaaga ctgggaattc ctacactgga 480
 attgggtgctg tgtaccctgc aaaccaagcc aagaaccagt cagatgaggc aaaagaggct 540
 aaggggagtg ggtatgagaa attagggccc agtcgagacc cagatccact gaacatctgt 600
 gtctttatct tgctgcttgt atttattgta gtcaaagtct ttacatcaga atgatgaaa 660
 taggcttgcc actttctctt attttaattc catggtagtc aatgaactgg ctgccacttt 720
 aatataactg aaaattcatt ttgagaccaa gcaggatcaa gttttagtaa taaacactgg 780
 tttcctagcc atcctctgaa aacagtatga aacatgacca agtacataat ggatttagta 840
 ataaatattg tcgaattgct 860

<210> 340
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig51660

<400> 340
 gctgcttgta tttattgtag tcaaagtgtt tacatcagaa tgatgaaaat aggccttgcca 60

<210> 341
 <211> 608
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig52490

<400> 341
 atcgtggcta gcggacagac acgagcctct tgggaatacc ttgtccatca cgtcatggcc 60

```

atgggtgcct tcttctccgg catcttttgg agcagctttg tcggtggggg tgtcttaaca 120
ctactgggtg aagtcagcaa catcttcctc accattcgca tgatgatgaa aatcagtaat 180
gccaggatc atctcctcta ccgggttaac aagtatgtga acctggatcat gtactttctc 240
ttccgcctgg cccctcaggc ctacctcacc catttcttct tgcgttatgt gaaccagagg 300
acctgggga ccttctgtgt gggatatcctg ctcatgctgg acgtgatgat cataatctac 360
ttttcccgcc tcttcgctc tgacttgagaa ctgagtggagg ggcacagagc ctggga caac aaaaacggac 480
aaggccagaa acagcttcat atggacactg ggacttagcc ccaagc ctgg gtgtcctctg 540
aggccagcct ctccaccttc tgagcctgcg cccacactat tgaaaa cact aatgaaagta 600
ctcctctg 608

```

<210> 342

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig52490

<400> 342

```

ccaggatcat ctctctacc ggggttaacaa gtatgtgaac ctggatcatgt actttctctt 60

```

<210> 343

<211> 1282

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 343

```

catgccagca cctttgaacc ggtctcttag aagaagacac acatcctggg tgtacagtgg 60
tgaaatgggg agtgggtgcc cattctgaaa aacgaggcat tctgtctcat tccctctgct 120
tagctgggtg gcaggggaga gagggaaatg ccaaaaactt ggagtgaagg atgatgctat 180
ttttttatatt taaatatatc ttccaggttat tttcttactg ttgcttcaga tctaattgtaa 240
aaggcagatg tccccctctc tccacccccg acgctgaccc cggcctcagt caccgctctt 300
tgcatgatca cagttctgtg ttctggcctg tggcagggcc ggggaagggcc gctggcttcc 360
gaacagacgt ggttgcctct cacgaggcgc atggggagcc cgcgggcctt aagctttgtc 420
gcagatgtca tcattggcag aattacttgt cttgaaaaat aagtagcatt gctgaaacac 480
acaaccgaat tctctacgat ggccatttgc tcattgtctt tctctgtgtg gtagtgagtg 540
accctggcag tgtttgccctg ctccagagtgg cccctcagaa caacaggggct ggccttggaa 600
aaa.ccccaaa acaggactgt ggtgacaact ctggtcaggt gtgatttgac atgagggccg 660
gagggcgggtg ctgacggcag gactggagag gctgcgtgcc cggcactggc agcgaggctc 720
gtgtgtcccc caggcagatc tgggcacttt cccaaccag gtttatgctg ctccagggaa 780
gcctcgggtg cagagtgggt ggcagatctg accatcccca cagacca.gaa acaagggaatt 840
tctgggatta cccagtcccc cttcaaccca gttgatgtaa ccacctcatt ttttacaaat 900
acagaatcta ttctactcag gctatgggcc tcgtcctcac tcagtta.ttg cgagtgttgc 960
tgtccgcatg ctccggggcc cacgtggctc ctgtgctcta gatcatggtg actccccgc 1020
cctgtgggtg gaatcgatgc cacggattgc aggccaaatt tcagatcgtg tttccaaaca 1080
ccc.ttgctgt gccctttaat gggattgaaa gcacttttac cacatggaga aatatatttt 1140
taattttgtga tgcttttcta caaggctcac tatttctgag tttaatgtgt ttccaacact 1200
taaggagact ctaatgaaag ctgatgaatt ttcttttctg tccaaac.aag taaaataaaa 1260
ataaaagtct atttagatgt tg 1282

```

<210> 344

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 344
ccactatatttc tgagtttaaat gtgttttccaa cacttaagga gactctaataa aaagctgatg 60

<210> 345
<211> 601
<212> DNA
<213> Homo sapiens

<300>
<308> Contig53641

<400> 345
tggaggctgt ggatgatgct ttcaagacaa tggatgtgga tatggccgag gaacatgcc 60
gggcccagat gagggccag atgaatatcg gggatgaagc gctgattgga cgggtggagct 120
gggatgacat acaagtcgag ctccctgacct gggatgagga cggagatttt ggcgatgcct 180
gggcccaggat cccctttgct ttctgggcca gataccatca gtacattctg aatagcaacc 240
gtgccaacag gagggccacg tggagagctg gcgtcagcag tggcaccaat ggaggggcca 300
gcaccagcgt cctagatggc cccagcacca gctccaccat ccggaccaga aatgctgcc 360
gagctggcgc cagcttcttc tcctggatcc agcacggtg acgaactgca gcgatcttac 420
tggccaagcc agagcgcctc ctctcagatt ccttcctgac acagcacctc aggcggcttc 480
ttcctgtcag tcggagggtg catgcaagat gaagctctct ttgctcttcc tgctttcatt 540
ttgtgctttt ccttgtgttt tcatgttttg ggtatcagtg ttacattaaa gttgcaaaaat 600
t 601

<210> 346
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig53641

<400> 346
ctttcatttt gtgcttttcc ttgtgttttc atgttttggg tatcagtgtt acattaaagt 60

<210> 347
<211> 751
<212> DNA
<213> Homo sapiens

<300>
<308> Contig54242

<400> 347
aattactcaa agaaggagcc atttcagtta actcaagtga atgaaagact tttggaatct 60
gcagtgggtc cttccctgtt gaccatttgg taacttgtaa tctgaccaa aactcttgag 120
ctgcaacagg ccttgccaga gggctcagga tgggaaggga agaaggggat aggaaaagaa 180
gaggtaattt tacatttccc ctttaaagta aatttttagcc aactcatcat tctgaaatgt 240
ccctataaag aatgagtcga actagaccag aagccagcct actccttctt acatagcttc 300
tccaacaggg gtagcaatga cctgtccact tcaaacacag ataaggcctg ccattcctcat 360
tgggttaaagg cacacgtgag actttcagtg ggctctgctg agaaggaagg cagcccagga 420
gtcagggtatg caggcattgc attgtcagtg tctgctctca gagtttacac attcaattgc 480
ttccaagggg gaatctcctg ctctgtgaat gctatcagac cccaaaggcc aaccttgggc 540
tgggtctatg tacgttcttc cgaagcactg atgatcaaaa ttgaagacac attcagaggt 600
ttgattgggt gagattaact ggtgtggtgg ttggtgtatg tatgttttat ttttatgtct 660
ttgtatgtag ttctacataa tgcaaatgt gcttcttgat ggacaagacc tcataactgt 720
gattaatatc aataaaaagg ggatgttgtg g 751

<210> 348

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54242

<400> 348
 gtaaattttta gccaaactcat catctctgaaa tgtccctata aagaatgagt cgaactagac 60

<210> 349
 <211> 637
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54661

<400> 349
 ggcagtgatg tctatgttga gat taactta tgtattgagg aaaatttgaa gtttattttt 60
 tcgatgaata aggctgtcaa atgatttagt atagattaat gacatctttt ttagaaatat 120
 taaagtggagt attcctcatt atgtcatcat ttctgataat tagagtgcta atttgaatgt 180
 tagataatgt ttccacatct ata cctatctt ctttctaggg cacttctgac cctgggggctt 240
 ggggatggcc tttaggccac agt agtgtct gtgttaagtt cactaaatgt gtatttaagt 300
 agaaacattc ctatgtaaaa atgtgtgtat gtgaacgtat gcatacattt ttattgtgca 360
 cctgtacatt gtgaagaagt agt ttggaaa ttgttaaagc acaaaccata aaagagtgtg 420
 gagttattaa atgatgtagc acaaatgtaa tgttttagctt ataaaaggct ctttctattt 480
 tctatggcaa agactttgac act tgaaaaa taaaaccaat atttgattta tttttgtaag 540
 tatttaggat attattttta ataatgatt gtccattatc aatataatag ttgtgaaatg 600
 atttaagtaa ataaacttta tgc ttctgtg tctgttg 637

<210> 350
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54661

<400> 350
 ctgtacattg tgaagaagta gtt tggaaat ttgtaaagca caaaccataa aagagtgtgg 60

<210> 351
 <211> 924
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55188

<400> 351
 gcgacaagta ccgcaagcgg gcactcatcc tgggtgtcact gctggccttt gccggcctct 60
 tcgccgccct cgtgctgtgg ctgtacatct acccattaa ctggccctgg atcgagcacc 120
 tcacctgctt ccccttcacc agc cgtctct gcgagaagta tgagctggac cagggtgctgc 180
 actgaccgct gggccacacg gctgcccctc agccctgctg gaacagggctc tgcctgcgag 240
 ggctgcccctc tgcagagcgc tctctgtgtg ccagagagcc agagacccaa gacagggccc 300
 gggctctgga cctgggtgcc cccctgcccag gcgaggctga ctccgcgtga gatggttggt 360
 taaggcgggg tttttctggg gcgtgaggcc tgtgagatcc tgacccaagc tcaggcacac 420
 ccaaggcacc tgccctctctg agtcttgggt ctcagttcct aatatcccgc tccttgctga 480
 gaccatctcc tggggcaggg tccctttctt cccaggtcct cagcgtgcc tctgctggtg 540

```

ccttctcccc ca.ctactact ggagcgtgcc cttgctgggg acgtggctgt gcc.ctcagtt 600
gccccagggt ct.gggtgccc accatgcccc ttctctcttc tcctcctacc tct.gccctgt 660
gagcccatcc at.aaggctct cagatgggac attgtgggaa aggctttggc cat.ggtctgg 720
gggcagagaa ca.agggggga gacacaagta gacctcaggt agaacgacac tgg.gcgaggc 780
caccctcagg cc.tgctccca gggagtgtct gaggcgcata aggcccgttt ttt.accagtt 840
tatatcacgg tc.ttcatttt taaaagtaac gctaactttg tacggacgat gtc.tcatgga 900
ttaaataata tt.ctttatgg cagt 924

```

<210> 352

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig5 5188

<400> 352

```

agtaacgcta ac.tttgtacg gacgatgtct catggattaa ataatatctt tta.tggcagt 60

```

<210> 353

<211> 699

<212> DNA

<213> Homo sapiens

<300>

<308> Contig5 5353

<400> 353

```

tgattatgcc aa.gagctcta aacagaagtt tgagaaggta aaaattaagt tgt.agtatct 60
gagttgtttt ta.ttttcttc ctttgggtgt tatgaaggta ttcataagaa ctt.taatttc 120
aggggaaaaa at.gcctgatt tgctattttt gacatttcct cgtctcttaa ga.agtcagtt 180
aaatatgttt tcatagttta tattcctggt tcatagatta ctgtgaaaca tgt.attttaa 240
cctatgaatt at.aaaatagt atttagattc tagcgtgagt taaatagatt agt.catatat 300
cttttagatt tgt.tggatttg acatgtaaat tatgtgttgt gtataagtaa gtt.agttact 360
aaacatatgg ca.tggttatt gataaacttg ttgctatttt tttccaaatg cta.tcagtgt 420
ttgtggactt tt.aaaaatta gtttgaattt tggaatgttc tgtgataaaa tata.aatttca 480
actattttgt ac.attttaa atgccatggt gtatatgtct gtatttaaaa atg.ttgtaaa 540
tatctgcatt tt.aagaatta tgaaagattt tcctcaaaaa tgacagaact ctc.catactt 600
aattgtgaca ca.ttataaga tatctgattt taagcttttg gattttgttc taaa.aattaa 660
gtttaaacat gc.tgaaaatt ccataaaaat aaaattttg 699

```

<210> 354

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig5 5353

<400> 354

```

taaaaatagta tt.tagattct agcgtgagtt aaatagatta gtcatatatc ttt.tagattt 60

```

<210> 355

<211> 809

<212> DNA

<213> Homo sapiens

<300>

<308> Contig5 6503

<400> 355

```

gcatgtgaga tgagtgactg ccggtgaatg tgtccacagt tga gaggttg gagcaggatg 60
agggaa tcct gtcaccatca ataactactt gtggagcgcc act ctgcccc agacgccacc 120
tgggcg gaca gcatggagct ctccatggcc aggctgcctg tgt gcatgtt ccctgtctgg 180
tgcccc tttg cccgcctcct gcaaacctca cagggtcccc aca caacagt gccctccaga 240
agcagc cctt cggaggcaga ggaaggaaaa tggggatggc tgggggctctc tccatcctcc 300
ttttct cctt gccttcgcat ggctggcctt cccctccaaa acc tccattc ccctgtctgcc 360
agcccc tttg ccatagcctg attttgggga ggaggaagg gcgatttgag ggagaagggg 420
agaaag cttt tggctgggtc tggtttcttc ccttcccaga ggg tcttact gttccagggt 480
ggcccc aggg caggcagggg ccacactatg cctgcgcctt ggt aaagggtg acccctgcca 540
tttacc agca gccctggcat gttcctgccc cacaggaata gaa tggaggg agctccagaa 600
actttc catc ccaaaggcag tctccgtggg tgaagcagac tggatttttg ctctgccctc 660
gacccc tttg cctcttttga gggaggggag ctatgctagg act ccaacct cagggactcg 720
ggtggc ctgc gctagcttct tttgatactg aaaactttta agg tgggagg gtggcaaggg 780
atgtgc ttaa taaatcaatt ccaagcctc 809

```

<210> 356

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56503

<400> 356

```

gaaaactttt aagggtgggag ggtggcaagg gatgtgctta ataaatcaat tccaagcctc 60

```

<210> 357

<211> 976

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56678

<400> 357

```

gaaggatata ctttgttata acttattatt ttgttctctg taaatacaag atgtttatag 60
gaaatatgta ttctgaactc tatctgcaga atgagtcact aca ccaaaat agttctatta 120
tttagaatgt gttaatttta aagggaacctg ataggatattt att tacatat gcgatccaca 180
tttgtgtgaa agcatgtgat cataactaacc cagcctcctg gaatgtcgcgt gtacgatgat 240
tgatgtcttt ttctcagtcct atagttacaa ttgttttagta tgctaatacag tccagttccc 300
tgagggtttaa gatcaaatat aaattactct gcttttcgac tca ttcagggt agcattgtac 360
ctgaacctga ttgctacttt ttcatcttaa atattatatt tcctcatcta atctgccttc 420
ccctcatcca cagacatttg gagaaggaaa tgggaggggtg tctgttatcc ctttctcttt 480
gctttgtccc cgttgttaga ctggcagcgt cagttgctcg gtgggcttgg ttagagccgt 540
gggtgaggca ggtggctggc ggggacaggg agaggctgag aggg aaagtgg tggcatttac 600
tgctctgaca cttccactgt cctgctggg gatgctgggg ccaaggcctg tggggcctgt 660
gaactgcaca gccaggagca aggaaccac taaatactcc gtcacctcca tgtcccctct 720
acagtgttaa attattacat aagcagggtga aaggtagaag gcgaattatg tgagtaaata 780
tggtctgttt tctcttcagc aaaaatgact atttttgtgt gtgactaatt tatttttatt 840
attgtaaaga tacaataaac cggttgaaat atctgctttg ttgacaagcg tgtgctttct 900
ctggccttat tcgcgttctg ttctctgca aatagcgcct tctaaaaaga agagtcagac 960
aataaaactgg ttgaaa 976

```

<210> 358

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig56678

<400> 358
tattacataa gcaggtgaaa ggtagaaggc gaattatgtg agtaaataatg gtctgttttc 60

<210> 359
<211> 1118
<212> DNA
<213> Homo sapiens

<300>
<308> Contig57584

<400> 359
agctgtttgtg catccagagg tgggaattggg gcccggcatt ccctcctcgt cccgggctgg 60
cccttgcccc caccctgcaa ctcttggttg agatgggctc agccaagagc gtcccagtc 120
caccagcgcg gcctccgccc acaacaagca tctggctcga gtggcggacc cccgttcacc 180
tagtgctggc atcctgcgca ctcccatcca ggtggagagc tctccacagc caggcctacc 240
agcaggggag caactggagg gtcttaaaca tgcccaggac tcagatcccc gctctcctac 300
tcttggtatt gcacggacac ctatgaagac cagcagtga gacccccaa gccactgggt 360

gaaacagctg agtgaagtat ttgaaactga agactctaaa tcaaactctc cccagagacc 420
tgttctgccc ccagaggcac ctttatcttc tgaattggac ttgcctctgg gtaccagtt 480
atctgttgag gaacagatgc caccttgga ccagactgag ttcccctcca aacagggtgt 540
ttccaaggag gaagcaagac agcccacaga aacccctgtg gccagccaga gctccgacaa 600
gccctcaagg gaccctgaga ctcccagatc ttcaggttct atgcgcaata gatggaaacc 660
aaacagcagc aaggtactag ggagatcccc cctcaccatc ctgcaggatg acaactcccc 720
tggcaccctg acactacgac agggtaagcg gccttcaccc ctaagtgaat atgttagtga 780
actaaaggaa ggagccattc ttggaactgg acgacttctg aaaactggag gacgagcatg 840
ggagcaaggc caggaccatg acaaggaaaa tcagcacttt cccttggtgg agagctaggc 900
cctgcatggc cccagcaatg cagtcaccca gggcctgggtg atatctgtgt cctctcacc 960
cttctttccc agggatactg aggaatggct tgttttctta gactcctcct cagctaccaa 1020
actgggactc acagctttat tgggctttct ttgtgtcttg tgtgtttctt ttatatataa 1080
ggaagtaatt ttaaatgtta ctttaaaaag gtatatgt 1118

<210> 360
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig57584

<400> 360
aggaatggct tgttttctta gactcctcct cagctaccaa actgggactc acagctttat 60

<210> 361
<211> 859
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63649

<400> 361
gtcgcagggt accagtgtgc ggagttcctg ttgccaaagc gaaggtggcc ctgggcaggc 60
acaggtgtgg tcatatcttc agccaacagg accatcctcc ggagggccac ctctggggac 120
ttcctacggg aagagagtga cagatttggt gcttctgtgt gtttctgccc cttcagtggt 180
gccgctgcgg gagacagcgg gtggatcctc cagcagcctg tctgctgagc ctgccttctc 240
aagtctactg ttaaaatcag gaccgggtcg tgtccgagcc tacaggccct gtctccgctc 300
cccaggcctg caggagttag gggctgcacc tgctcgtgtg agaggagag gcagatttag 360
tggacgcctg gcatggactc ggactggcct ttggaagctc cctgccctga cgggttgcc 420
gtcaccactg cgaagtggag cttggaggac ctgcacctga gaaaggctgt gtgtgtgtct 480

```

gggtccacac ctgccagagc taacttactg ccagacggcg acttactgtg ggccaccctc 540
agtgaaccgg ggtgtcctca gctggcccta cagagcactt ctgtgctggg gatgagtagg 600
aactctgggc gaggagggtc ccagcgccgc ccctcgatac agccctgctc tgccctctgc 660
ccgtacttat accaggtggg atccctgccc tgcattgcct ggggattggc tgggcttggg 720
cacgccctgc tgtggaactg gatgttttca gggagcccag cctttcctca tgtcaacaca 780
gttcacaata tagttttcaa agtacagttt aaaactcaaa agtaaacttt tcagcaactc 840
aaaaaaaaaa aaaaaaaaaa 859

```

<210> 362
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63649

<400> 362
 cagcctttcc tcatgtcaac acagttcaca atatagtttt caaagtacag tttaaaactc 60

<210> 363
 <211> 1170
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63525

<400> 363
 gccatggctc cctggggcga gcgagcactc gcggctgaac ccgctgcgcg cgggtgtggct 60
 cacgctgacc gccgccttcc tgctgaccct actgctgcag ctccctgccgc cgggcctgct 120
 cccgggctgc gcgatcttcc aggacctgat ccgctatggg aaaaccaagt gtggggagcc 180
 gtcgcgcccc gccgcctgcc gagcctttga tgtccccaag agatattttt cccactttta 240
 tatcatctca gtgctgtgga atggcttcct gctttggtgc cttactcaat ctctgttctc 300
 gggagcacct tttccaagct ggcttcatgg tttgctcaga attctcgggg cggcacagt 360
 ccaggagggg gagctggcac tgtctgcatt cttagtgtca gtatttctgt ggctgcacag 420
 cttacgaaga ctcttcgagt gcctctacgt cagtgtcttc tccaatgtca tgattcacgt 480
 cgtgcagtac tgttttggac ttgtctatta tgtccttgtt ggctaactg tgctgagcca 540
 agtgccaatg gatggcagga atgctacata acagggaata atctattgat gcaagcacgg 600
 tggttccata ttcttgggat gatgatgttc atctggatcat ctgcccatca gtataagtgc 660
 catgttattc tcggcaatct caggaaaaat aaagcaggag tggtcattca ctgtaaccac 720
 aggatcccat ttggagactg gtttgaatat gtttcttccc ctaactactt agcagagctg 780
 atgatctacg tttccatggc cgtcaccttt gggttccaca acttaacttg gtggctagt 840
 gtgacaaatg tcttctttaa tcaggccctg tctgcctttc tcagccacca attctacaaa 900
 agcaaatttg tctcttacc gaagcatagg aaagctttcc taccattttt gttttaagtt 960
 aacctcagtc atgaagaatg caaaccaggt gatggtttca atgcctaagg acagtgaagt 1020
 ctggagccca aagtacagtt tcagcaaagc tgtttgaaac tctccattcc atttctatac 1080
 cccacaagtt ttactgaat gagcatgcag tgccactcaa gaaaatgaat ctccaaagta 1140
 tcttcaaaga attaattact aatggcagat 1170

<210> 364
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63525

<400> 364
 ctcttaccgg aagcatagga aagctttcct accatttttg ttttaagtta acctcagtc 60

<210> 365

<211> 632
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 365
 aagaatgcta agatgatttc agatatcgaa aagaaaaggc agcgtatgat tgaagtccag 60
 gatgaactgc ttccggttaga gccacagctg aaacaactac aaacaaaata tgatgaactt 120
 aaagagagaa agtcttccct taggaatgca gcataatttct tatctaattt aaaacagctt 180
 tatcaagatt attcagatgt tcaagctcaa gaaccaaacy taaaggaaac gtatgattca 240
 tccagccttc cagctctgtt atttaaagca agaacacttc tgggagccga aagccatctg 300
 cgaaatatca accatcagtt agagaagctc cttgaccagg gatgagaaga gcagtctact 360
 aaaatgtgcc tataggaaga ctagtctcat gctgttacct tctgaaactg tacctttata 420
 aatcaattgt tttgcaaaga agttatggcc tacttagaat ctaaaatttg ttattcaaat 480
 taaatggctg tgaacaatgt taaatagcat cagtttgtcc aatagtttta aaggccataa 540
 tcatcttttc tgggttaatat cttgagtaat tttaaaatgt tgacacctta atcgggtcca 600
 ggtatgagcc ataataaact tgtaaaatta ag 632

<210> 366
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 366
 ggctgtgaac aatgtttaa atgcacagtt tgtccaatag ttttaaaggc cataatcatc 60

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US05/07894

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68

US CL : 435/6

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST, STN, brca1, correlat?, tumor?, tumour?, cancer?, carcinoma

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Van't Veer et al, Gene expression profiling predicts clinical outcome of breast cancer, Nature, 31 January 2002, Vol. 415, pages 530-536, see entire document.	1-11 and 33-41
X	WO 02/103320 A2 (ROSETTA INPHARMATICS INC.) 27 December 2002 (27.12.2002), see especially page 1, line 26 through page 10, line 34.	1-11 and 33-41

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

24 September 2005 (24.09.2005)

Date of mailing of the international search report

23 NOV 2005

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US

Commissioner for Patents

P.O. Box 1450

Alexandria, Virginia 22313-1450

Facsimile No. (571) 273-3201

Authorized officer

James Martinell

Telephone No. (571) 272-0719

INTERNATIONAL SEARCH REPORT

International application No. _____

PCT/US05/07894

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.: _____
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 12-32 and 42-87
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 12-32 and 42-87 refer to one or more Tables in the specification and so do not comply with PCT Rule 6.2(a).

3. ☐ Claims Nos.: _____
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of any additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.